A new method for estimating linear models with time varying parameters and errors in the observed data

Willemstein, A.P.

Publication date:
1978

Document Version
Publisher's PDF, also known as Version of record

A NEW METHOD FOR ESTIMATING LINEAR MODELS WITH TIME VARYING PARAMETERS AND ERRORS IN THE OBSERVED DATA

by

A. P. WILLEMSTEIN

Research memorandum

TILBURG UNIVERSITY
DEPARTMENT OF ECONOMICS
Postbus 90135 - 5000 LE Tilburg
Netherlands

FEW
72

A NEW METHOD FOR ESTIMATING LINEAR MODELS WITH TIME VARYING PARAMETERS AND ERRORS IN THE OBSERVED DATA

A.P. WILLEMSTEIN

R11

T linear models
T estimation

Research Memorandum

Tilburg University
Department of Econometrics
P.O.B. 90135-5000 LE Tilburg
The Netherlands

The research reported here has been sponsored during the period sept. 1976 - dec. 1977 by the Netherlands Organization for the Advancement of Pure Research (Z.W.O.)

Contents

Page

Abstract

This paper deals with an identification method for the unknown parameters in a linear discrete dynamic econometric system.
The identification method can be formulated as an approximation problem and distincts itself from conventional methods by charging observation errors on the endogenous and exogenous variables and by introducing small time varying fluctuations on the (autonomous) parameters.
Some experimental results will be given for L.R. Kleins Model I.
Finally the system will be considered as a stochastic model and some statistical properties of the identification (estimation) method will be discussed.

Notation

The standard inner product of the vectors x and y	: $x'y$
The euclidean norm of a vector x	: $\|x\| = (x'x)^{\frac{1}{2}}$
The i^{th} component of the vector x	: $(x)_i$
The set of nxm matrices	: $M_{n,m}$
The transposed of a matrix A	: A'
The euclidean norm of a matrix A	: $\|A\|$
The nxn identity matrix	: $I_n \in M_{n,n}$
The i,j^{th} element of the matrix A	: A_{ij}
The i^{th} column of the matrix A	: $(A)_i$
The set of positive definite matrices	: PD
The set of positive semi definite matrices	: PSD
The gradient of a vector function f(x)	: $\nabla_x f$
The Hessian of a vector function f(x)	: $\nabla_{xx} f$
The expectation of a stochastic vector (variable) $\underline{x}$	: $E\,\underline{x}$
The variance of a stochastic variable $\underline{y}$	: $var(\underline{y})$
The variance-covariance matrix of a stochastic vector $\underline{x} \in \mathbb{R}^p$	: $VAR(\underline{x}) \in M_{p,p}$

The covariance of two stochastic variables
$\underline{y}_1$ and $\underline{y}_2$: $\mathrm{cov}(\underline{y}_1, \underline{y}_2)$

The covariance matrix of two stochastic vectors
$\underline{x}_1 \in \mathbb{R}^p$ and $\underline{x}_2 \in \mathbb{R}^q$: $\mathrm{COV}(\underline{x}_1, \underline{x}_2) \in M_{p,q}$

1. Introduction

We usually consider econometric models with the structural form:

$$y_{t+1} = Ay_t + Bx_t \quad (t=0,1,\dots.) \tag{1}$$

where the vector $y_t \in \mathbb{R}^n$ represents the set of endogenous variables and the vector $x_t \in \mathbb{R}^m$ represents the set of exogenous variables in time t.
The matrices $A \in M_{n,n}$ and $B \in M_{n,m}$ contain the unknown structural parameters of the system.
Equation (1) never gives an exact description of the econometric system. The following influences are neglected, though they may be relevant to the system as well:

(i) Nonlinear terms in x_t and y_t
(ii) Another lagging structure
(iii) More relevant (exogenous) variables
(iv) Observation errors in the data
(v) Small time dependent disturbances on the autonomous parameters.

Several methods are available to evaluate the unknown matrices A and B on the basis of observed (measured) values of the endogenous and exogenous variables during a certain time period, say $[0,1,\dots,T]$ (e.g. see [1], [2]). One of the most widely applied methods is the method of ordinary least squares (OLS). This method is based on the assumption that all uncertain unfluences, described above, are fully attributed to a residual r_t ($r_t \in \mathbb{R}^n$):

$$y_{t+1} = Ay_t + Bx_t + r_t \quad (t=0,1,\dots,T-1) . \tag{2}$$

The OLS method consists of minimization of the following function:

$$\min \{ \sum_{t=0}^{T-1} \|r_t\|^2 | y_{t+1} = Ay_t + Bx_t + r_t,\ r_t \in \mathbb{R}^n \quad (t=0,\ldots,T-1),$$

$$A \in M_{n,n},\ B \in M_{n,m} \}. \qquad (3)$$

So the sum of the squared norms of the T residuals is to be minimized. Related with the OLS method is the so-called weighted least squares method (WLS). Given the matrices $Q_t \in M_{n,n} \cap PD$, Q_t diagonal $(t=0,\ldots,T-1)$, we minimize the function:

$$\min \{ \sum_{t=0}^{T-1} r_t' \, Q_t \, r_t | y_{t+1} = Ay_t + Bx_t + r_t,\ r_t \in \mathbb{R}^n \ (t=0,\ldots,T-1)$$

$$A \in M_{n,n},\ B \in M_{n,m} \}. \qquad (4)$$

The WLS method gives the possibility of weighting the residuals according to some criterion. In the case when the OLS residuals are big in norm, it is acceptable trying to attribute those unknown influences partly to small observation errors in the data on one side and to small time varying fluctuations on the parameters on the other side:

$$y_{y+1} + \xi_{t+1} = (A+E_t)(y_t+\xi_t) + (B+F_t)(x_t+\eta_t) + r_t$$

$$(t=0,\ldots,T-1). \qquad (5)$$

Here the vectors $\xi_t \in \mathbb{R}^n$ $(t=0,\ldots,T)$ and $\eta_t \in \mathbb{R}^m$ $(t=0,\ldots.,T-1)$ are corresponding to the observation errors in y_t and x_t respectively. The matrices $E_t \in M_{n,n}$ and $F_t \in M_{n,m}$ $(t=0,\ldots,T-1)$ represent the fluctuations on A and B respectively. In argreement with the OLS method we introduce the following criterion for evaluating the matrices A and B:

$$\min\{ \sum_{t=0}^{T-1} [\|r_t\|^2 + \|\xi_t\|^2 + \|\eta_t\|^2 + \|E_t\|^2 + \|F_t\|^2 + \|\xi_T\|^2 |$$

$$|y_{t+1} + \xi_{t+1} = (A+E_t)(y_t+\xi_t) + (B+F_t)(x_t+\eta_t) + r_t,$$

$$A,E_t \in M_{n,n},\ B,F_t \in M_{n,m},\ \xi_t,\xi_T,r_t \in \mathbb{R}^n,$$

$$\eta_t \in \mathbb{R}^m\ (t=0,....,T-1)\}. \qquad (6)$$

We notice that this method can be considered as an extension of the OLS method. We shall call this method the composed ordinary least squares method (COLS).
The following criterion relates COLS to the WLS method:

$$\min\{ \sum_{t=0}^{T-1} [r_t' Q_t r_t + \xi_t' R_t \xi_t + \eta_t' P_t \eta_t + \sum_{i=1}^{n} (E_t)_i' S_{ti} (E_t)_i +$$

$$+ \sum_{j=1}^{m} (F_t)_j' T_{tj} (F_t)_j] + \xi_T' R_T \xi_T \ |y_{t+1} + \xi_{t+1} =$$

$$= (A+E_t)(y_t+\xi_t) + (B+F_t)(x_t+\eta_t) + r_t,$$

$$A,E_t \in M_{n,n},\ B,F_t \in M_{n,m},\ \xi_t,\xi_T,\ r_t \in \mathbb{R}^n,$$

$$\eta_t \in \mathbb{R}^m\ \ (t=0,....,T-1)\}. \qquad (7)$$

Here we assume that the matrices $Q_t \in M_{n,n}$, $R_t \in M_{n,n}$, $P_t \in M_{m,m}$, $S_{ti} \in M_{n,n}$ and $T_{tj} \in M_{n,n}$ are known, positive definite and diagonal. Criterion (7) presents the possibility of weighting e.g. the contribution of ξ_t with regard to the contribution of E_t. Furthermore it is possible in (7) to replace absolute errors by relative errors. We call (7) the composedweighted least squares method (CWLS).
In chapter 2 we derive an iterative procedure for solving the COLS problem (6). If is easy to verify that the CWLS problem can be treated in the same way. In chapter 3 we assume that the residuals

r_t are all equal to zero and the minimization is with respect to A, B, ξ_t, η_t, E_t and F_t only.
Current econometric models (e.g. Models I and III of L.R. Klein, Klein-Goldberger Model) are not given in the structural form (1). Those models consist of two subsets of equations: a set of reaction equations, where all the unknown parameters appear, and a set of definition equations (identities), which hold exactly for all the observations. In chapter 4 we introduce a general iterative method for estimating this kind of models, by extending the method of chapter 2.
In chapter 5 we give some experimental results. The Model I of L.R. Klein (see [3]) is choosen as an example. It appears that a very significant reduction of the OLS residuals is achieved by the introduction of comparatively small fluctuations on the data and the parameters.
In the following two chapters we consider <u>stochastic</u> models i.e. models where the state vector y_t, the residual r_t, the observation errors ξ_t, η_t and the time varying fluctuations E_t, F_t are random. We shall restrict ourselves to models with only stochastic residuals and stochastic parameter fluctuations. Observation errors are left out of consideration.
First we dicuss in chapter 6 the general static model, which leads to a surprising result. Namely, if the identification method of chapter 2 is considered as a statistical estimation method with respect to the unknown parameters than this method is equivalent to an estimation method of a heteroscedastic regression model.
Finally in chapter 7 some remarks are made concerning the stochastic dynamic model.

2. An iterative procedure solving the COLS problem.

From (6) it follows that the COLS problem can be stated in the form:

$$\min\{ \sum_{t=0}^{T-1} [\|y_{t+1} + \xi_{t+1} - (A+E_t)(y_t+\xi_t) - (B+F_t)(x_t+\eta_t)\|^2 +$$

$$+ \|\xi_t\|^2 + \|\eta_t\|^2 + \|E_t\|^2 + \|F_t\|^2] + \|\xi_T\|^2 \; | A,E_t \in M_{n,n},$$

$$B,F_t \in M_{n,m}, \; \xi_t, \; \xi_T \in \mathbb{R}^n, \; \eta_t \in \mathbb{R}^m \quad (t=0,\ldots,T-1)\}. \tag{8}$$

The residuals r_t $(0 \leq t \leq T-1)$ are eliminated; after solving (8) their optimal values are determined by

$$r_t = y_{t+1} + \xi_{t+1} - (A+E_t)(y_t+\xi_t) - (B+F_t)(x_t+\eta_t) \quad .$$

$$(t=0,\ldots,T-1) \tag{9}$$

The problem is now reduced to a minimization problem <u>without</u> constraints. The total number of unknown parameters is (T+1) $(n^2 + nm + n + m) - m$.

This number grows fast when the dimension of the problem and the lenght of the sample period increase. Not the intricacy of the function but the large number of the unknown parameters turn out to be the most difficult obstacle when solving (8).

Several numerical methods are available for solving unconstrained minimization problems (e.g. see [4], [5]). We mention here the conjugate gradient method, because the gradients can be calculated rather simple. Furthermore a Gauss-Newton method could be suitable, because the objective function is a sum of squares.

Hoewever, we shall develop here another procedure. The reason for it can be the best explained as follows. It is possible splitting up the original problem (8) into two subproblems. The subproblems are:

(i) The optimization in (8) is only with respect to $\xi_t \in \mathbb{R}^n$ $(t=0,\ldots,T)$ and $\eta_t \in \mathbb{R}^m$ $(t=0,\ldots,T-1)$. The values of the matrices A, B, E_t and F_t $(t=0,\ldots,T-1)$ are assumed to be fixed.

(ii) The optimization in (8) is only with respect to $A, E_t \in M_{n,n}$, $B, F_t \in M_{n,m}$ $(t=0,\ldots,T-1)$. The values of the vectors ξ_t $(t=0,\ldots,T)$ and η_t $(t=0,\ldots,T-1)$ are assumed to be fixed.

These subproblems have the favourable property that they can be solved partially in an analytical way and this means that we will exploit the structure of the problem.
The two subproblems are the basis for an iterative numerical method for solving the original problem. First of all we define the function $G:\mathbb{R}^{(T+1)(n^2+nm+n+m)-m} \to \mathbb{R}$ as follows:

$$G(A,B,\xi_0\ldots\xi_T, \eta_0,\ldots,\eta_{T-1}, E_0,\ldots,E_{T-1}, F_0,\ldots,F_{T-1}) := \sum_{t=0}^{T-1} [\|y_{t+1} + \xi_{t+1} - (A+E_t)(y_t+\xi_t) - (B+F_t)(x_t+\eta_t)\|^2 + \|\xi_t\|^2 + \|\eta_t\|^2 + \|E_t\|^2 + \|F_t\|^2] + \|\xi_T\|^2 \quad . \tag{10}$$

Then the problem is to minimize the function G with respect to $A \in M_{n,n}$, $B \in M_{n,m}$, $\xi_0,\ldots,\xi_T \in \mathbb{R}^n$, $\eta_0,\ldots, \eta_{T-1} \in \mathbb{R}^m$, $E_0,\ldots,E_{T-1} \in M_{n,n}$ and $F_0,\ldots,F_{T-1} \in M_{n,m}$.
We notice that if we define the function $H:\mathbb{R}^{n^2+nm} \to \mathbb{R}$ by

$$H(A,B) := G(A,B,0,\ldots,0,0,\ldots,0,0,\ldots,0,0,\ldots,0) \tag{11}$$

then the minimization problem $\min_{A,B} H(A,B)$ is identical as the OLS problem!

We shall now derive the necessary conditions for the function G having a minimum by calculating the gradients with respect to $A,B,\xi_0,\ldots,\xi_T,\eta_0,\ldots,\eta_{T-1}, E_0,\ldots,E_{T-1}, F_0,\ldots,F_{T-1}$ and putting them zero afterwards.

We notice that in addition to the usual gradient of a function with respect to a vector, we shall use here the gradient with respect to a matrix (see Appendix A). Furthermore the shorter notation $G(A,B,\underline{\xi},\underline{\eta},\underline{E},\underline{F})$ will be used.

The concerning gradients are:

$$\nabla_A\, G(A,B,\underline{\xi},\underline{\eta},\underline{E},\underline{F}) =$$

$$- 2 \sum_{t=0}^{T-1} [\, y_{t+1} + \xi_{t+1} - (A+E_t)(y_t+\xi_t) - (B+F_t)(x_t+\eta_t)]\,(y_t+\xi_t)' \qquad (12)$$

$$\nabla_B\, G(A,B,\underline{\xi},\underline{\eta},\underline{E},\underline{F}) =$$

$$- 2 \sum_{t=0}^{T-1} [\, y_{t+1} + \xi_{t+1} - (A+E_t)(y_t+\xi_t) - (B+F_t)(x_t+\eta_t)]\,(x_t+\eta_t)' \qquad (13)$$

$$\nabla_{\xi_0}\, G(A,B,\underline{\xi},\underline{\eta},\underline{E},\underline{F}) =$$

$$2\,\xi_0 - 2(A+E_0)'[\, y_1 + \xi_1 - (A+E_0)(y_0+\xi_0) - (B+F_0)(x_0 + \eta_0)] \qquad (14)$$

$$\nabla_{\xi_t}\, G(A,B,\underline{\xi},\underline{\eta},\underline{E},\underline{F}) =$$

$$2\,\xi_t - 2(A+E_t)'[\, y_{t+1} + \xi_{t+1} - (A+E_t)(y_t+\xi_t) - (B+F_t)(x_t+\eta_t)] +$$

$$+ 2[\, y_t + \xi_t - (A+E_{t-1})(y_{t-1}+\xi_{t-1}) - (B+F_{t-1})(x_{t-1}+\eta_{t-1})]$$

$$(t=1,\ldots,T-1) \qquad (15)$$

$$\nabla_{\xi_T} G(A,B,\underline{\xi},\underline{\eta},\underline{E},\underline{F}) =$$

$$2\ \xi_T + 2[y_t + \xi_T - (A+E_{T-1})(y_{T-1}+\xi_{T-1}) - (B+F_{T-1})(x_{T-1}+\eta_{T-1})]$$

$$\nabla_{\eta_T} G(A,B,\underline{\xi},\underline{\eta},\underline{E},\underline{F}) =$$

$$2\ \eta_t - 2(B+F_t)'[y_{t+1}+\xi_{t+1} - (A+E_t)(y_t+\xi_t) - (B+F_t)(x_t+\eta_t)]$$

$$(t=0,\ldots,T-1) \qquad (17)$$

$$\nabla_{E_t} G(A,B,\underline{\xi},\underline{\eta},\underline{E},\underline{F}) =$$

$$2\ E_t - 2[y_{t+1} + \xi_{t+1} - (A+E_t)(y_t+\xi_t) - (B+F_t)(x_t+\eta_t)](y_t+\xi)'$$

$$(t=0,\ldots,T-1) \qquad (18)$$

$$\nabla_{F_t} G(A,B,\underline{\xi},\underline{\eta},\underline{E},\underline{F}) =$$

$$2\ F_t - 2[y_{t+1} + \xi_{t+1} - (A+E_t)(y_t+\xi_t) - (B+F_t)(x_t+\eta_t)](x_t+\eta_t)'.$$

$$(t=0,\ldots\ldots,T-1) \qquad (19)$$

With (9) we get

$$\nabla_A G(A,B,\underline{\xi},\underline{\eta},\underline{E},\underline{F}) = -2 \sum_{t=0}^{T-1} r_t(y_t+\xi_t)' \qquad (20)$$

$$\nabla_B G(A,B,\underline{\xi},\underline{\eta},\underline{E},\underline{F}) = -2 \sum_{t=0}^{T-1} r_t(x_t+\eta_t)' \qquad (21)$$

$$\nabla_{\xi_0} G(A,B,\underline{\xi},\underline{\eta},\underline{E},\underline{F}) = 2\ \xi_0 - 2(A+E_0)'r_0 \qquad (22)$$

$$\nabla_{\xi_t} G(A,B,\underline{\xi},\underline{\eta},\underline{E},\underline{F}) = 2\ \xi_t - 2(A+E_t)'r_t + 2\ r_{t-1}$$

$$(t=1,\ldots\ldots,T-1) \qquad (23)$$

$$\nabla_{\xi_T} G(A,B,\underline{\xi},\underline{\eta},\underline{E},\underline{F}) = 2\ \xi_T + 2\ r_{T-1} \tag{24}$$

$$\nabla_{\eta_t} G(A,B,\underline{\xi},\underline{\eta},\underline{E},\underline{F}) = 2\ \eta_t - 2(B+F_t)'r_t \quad (t=0,\ldots,T-1) \tag{25}$$

$$\nabla_{E_t} G(A,B,\underline{\xi},\underline{\eta},\underline{E},\underline{F}) = 2\ E_t - 2\ r_t(y_t+\xi_t)' \quad (t=0,\ldots.,T-1) \tag{26}$$

$$\nabla_{F_t} G(A,B,\underline{\xi},\underline{\eta},\underline{E},\underline{F}) = 2\ F_t - 2\ r_t(x_t+\eta_t)' \quad (t=0,\ldots.,T-1). \tag{27}$$

Setting all these gradients equal to zero we get the first order conditions for a stationary point:

$$\boxed{\sum_{t=0}^{T-1} r_t(y_t+\xi_t)' = 0} \tag{28}$$

$$\boxed{\sum_{t=0}^{T-1} r_t(x_t+\eta_t)' = 0} \tag{29}$$

$$\xi_0 = (A+E_0)'r_0 \tag{30}$$

$$\xi_t = (A+E_t)'r_t - r_{t-1} \qquad (t=1,\ldots.,T-1) \tag{31}$$

$$\xi_T = -r_{T-1} \tag{32}$$

$$\boxed{\eta_t = (B+F_t)'r_t} \qquad (t=0,\ldots.,T-1) \tag{33}$$

$$\boxed{E_t = r_t(y_t+\xi_t)'} \qquad (t=0,\ldots,T-1) \tag{34}$$

$$\boxed{F_t = r_t(x_t+\eta_t)'} \qquad (t=0,\ldots,T-1) \quad . \tag{35}$$

If a point $(A,B,\underline{\xi},\underline{\eta},\underline{E},\underline{F})$ satisfies the conditions (28) .... (35), it is not necessary the minimum we are looking for. It is possible that this point is a saddle point or a local minimum.
However, a stationary point can never generate a (local) maximum because the partial Hessian

$$\nabla_{\eta_t \eta_t} G(A,B,\underline{\xi},\underline{\eta},\underline{E},\underline{F}) = 2\,[\,I_m + (B+F_t)'(B+F_t)] \tag{36}$$

is positive definite everywhere!

<u>Remark</u>. From (28), (29), (34) and (35) it follows that

$$\sum_{t=0}^{T-1} E_t = 0 \tag{37}$$

and

$$\sum_{t=0}^{T-1} F_t = 0 \quad . \tag{38}$$

We shall now discuss the two subproblems, we mentioned before. First we shall investigate the problem that the optimization takes place over only the ξ_t's $(t=0,\ldots,T)$ and the η_t's $(t=0,\ldots,T-1)$ assuming that the matrices $A,B,E_0,\ldots,E_{t-1}, F_0,\ldots,F_{T-1}$ are fixed.
Hence, assuming knowledge about the parameters and the fluctuations of the parameters, we attribute errors to the data in an optimal way.

Second we shall minimize the function G with respect tot A, B, $E_0,\dots E_{T-1}$, $F_0,\dots,F_{T-1}$ while the values of ξ_t and η_t are fixed. Now, assuming knowledge about the errors on the data, we attribute fluctuations of the parameters in an optimal way and calculate the optimal parameters.

I. Fixed A, B, $\underline{E}$ and $\underline{F}$.

We are only charging residuals r_t $(t=0,\dots,T-1)$ and observation errors ξ_t $(t=0,\dots,T)$ and η_t $(t=0,\dots,T-1)$. The first order conditions (30), (31), (32) and (33) play a role here. Define $A_t: = A+E_t$, $B_t: = B+F_t$, $s_t: = y_{t+1} - A_t\, y_t - B_t\, x_t$ $(t=0,\dots,T-1)$. Then the residuals r_t can be written as

$$r_t = s_t + \xi_{t+1} - A_t\, \xi_t - B_t\, \eta_t \quad (t=0,\dots,T-1). \tag{39}$$

The conditions (30), (31), (32) and (33) are respectively equal to

$$\begin{cases} \xi_0 = A_0'\, r_0 & (40) \\ \xi_t = A_t'\, r_t - r_{t-1} \quad (t=1,\dots,T-1) & (41) \\ \xi_T = -r_{T-1} & (42) \\ \eta_t = B_t'\, r_t \quad (t=0,\dots,T-1)\quad . & (43) \end{cases}$$

From (39) it follows that

$$\begin{cases} r_0 = s_0 + A_1'\, r_1 - r_0 - A_0\, A_0'\, r_0 - B_0\, B_0'\, r_0 & (44) \\ r_t = s_t + A_{t+1}'\, r_{t+1} - r_t - A_t\, A_t'\, r_t + A_t\, r_{t-1} - B_t\, B_t'\, r_t & \\ \qquad T=1,\dots,T-2) & (45) \end{cases}$$

$$\left\{ \begin{aligned} r_{T-1} &= s_{T-1} - r_{T-1} - A_{T-1} A'_{T-1} r_{T-1} + \\ &\quad + A_{T-1} r_{T-2} - B_{T-1} B'_{T-1} r_{T-1} \,. \end{aligned} \right. \tag{46}$$

Hence

$$\left\{ \begin{aligned} &(2 I_n + A_0 A'_0 + B_0 B'_0) r_0 - A'_1 r_1 = s_0 & (47) \\ &- A_t r_{t-1} + (2I_n + A_t A'_t + B_t B'_t) r_t - A'_{t+1} r_{t+1} = s_t & (48) \\ &\qquad (t=1,\ldots,T-2) & \\ &-A_{T-1} r_{T-2} + (2I_n + A_{T-1} A'_{T-1} + B_{T-1} B'_{T-1}) r_{T-1} = s_{T-1} \,. & (49) \end{aligned} \right.$$

These equations can be written in the following form:

$$\begin{bmatrix} 2I_n + A_0A'_0 + B_0B'_0 & -A'_1 & & & 0 \\ -A_1 & 2I_n + A_1A'_1 + B_1B'_1 & -A'_2 & & \\ & \ddots & \ddots & \ddots & \\ & -A_{T-2} & 2I_n + A_{T-2}A'_{T-2} + B_{T-2}B'_{T-2} & & -A'_{T-1} \\ 0 & & -A_{T-1} & & 2I_n + A_{T-1}A'_{T-1} + B_{T-1}B'_{T-1} \end{bmatrix} \begin{bmatrix} r_0 \\ \vdots \\ \vdots \\ r_{T-1} \end{bmatrix} = \begin{bmatrix} s_0 \\ \vdots \\ \vdots \\ s_{T-1} \end{bmatrix} . \tag{50}$$

The block tridiagonal matrix in (50) is non-singular (even positive definite); for this matrix can be written as

$$\begin{bmatrix} I_n+B_0B_0' & & & 0 \\ & I_n+B_1B_1' & & \\ & & \ddots & \\ 0 & & & I_n+B_{T-1}B_{T-1}' \end{bmatrix} + \left[\begin{array}{c|c|c} A_0 & -I_n & 0 \\ & A_1 & -I_n \\ & \ddots & \ddots \\ 0 & A_{T-1} & -I_n \end{array}\right]\left[\begin{array}{c|c|c} A_0 & -I_n & 0 \\ & A_1 & -I_n \\ & \ddots & \ddots \\ 0 & A_{T-1} & -I_n \end{array}\right]' . \tag{51}$$

From the equations (50), (40), (41), (42) en (43) we can determine the optimal $\underline{\xi}$, $\underline{\eta}$ and $\underline{r}$.

II. Fixed $\underline{\xi}$ and $\underline{\eta}$

We shall minimize the function G with respect to the structural matrices A, B and the parameter fluctuations E_t, F_t (t=0,...,T-1). The first order conditions (28), (29), (34) and (35) are of importance.
We define here $\tilde{y}_t := y_t + \xi_t$ (t=0,...,T) and $\tilde{x}_t := x_t + \eta_t$ (t=0,...,T-1). Then we have

$$r_t = \tilde{y}_{t+1} - (A+E_t)\tilde{y}_t - (B+F_t)\tilde{x}_t \quad (t=0,\dots,T-1) \tag{52}$$

The conditions (28), (29), (34) and (35) read respectively:

$$\sum_{t=0}^{T-1} r_t\, \tilde{y}_t' = 0 \tag{53}$$

$$\sum_{t=0}^{T-1} r_t\, \tilde{x}_t' = 0 \tag{54}$$

$$E_t = r_t\, \tilde{y}_t' \qquad (t=0,\dots,T-1) \tag{55}$$

$$F_t = r_t\, \tilde{x}_t' \qquad (t=0,\dots,T-1). \tag{56}$$

From (52), (55) and (56) it follows:

$$r_t = \tilde{y}_{t+1} - A\,\tilde{y}_t - B\,\tilde{x}_t - r_t\,[\|\tilde{y}_t\|^2 + \|\tilde{x}_t\|^2] \,.$$

$$(t=0,\ldots,T-1) \qquad (57)$$

Hence

$$r_t = \alpha_t^2\,(\tilde{y}_{t+1} - A\,\tilde{y}_t - B\,\tilde{x}_t) \quad (t=0,\ldots,T-1) \qquad (58)$$

where

$$\alpha_t := [\,1 + \|\tilde{y}_t\|^2 + \|\tilde{x}_t\|^2]^{-1/2} \ (t=0,\ldots,T-1) \,. \qquad (59)$$

Now the conditions (53) and (54) are similar to:

$$\sum_{t=0}^{T-1} \alpha_t^2(\tilde{y}_{t+1} - A\,\tilde{y}_t - B\,\tilde{x}_t)\tilde{y}_t' = 0 \qquad (60)$$

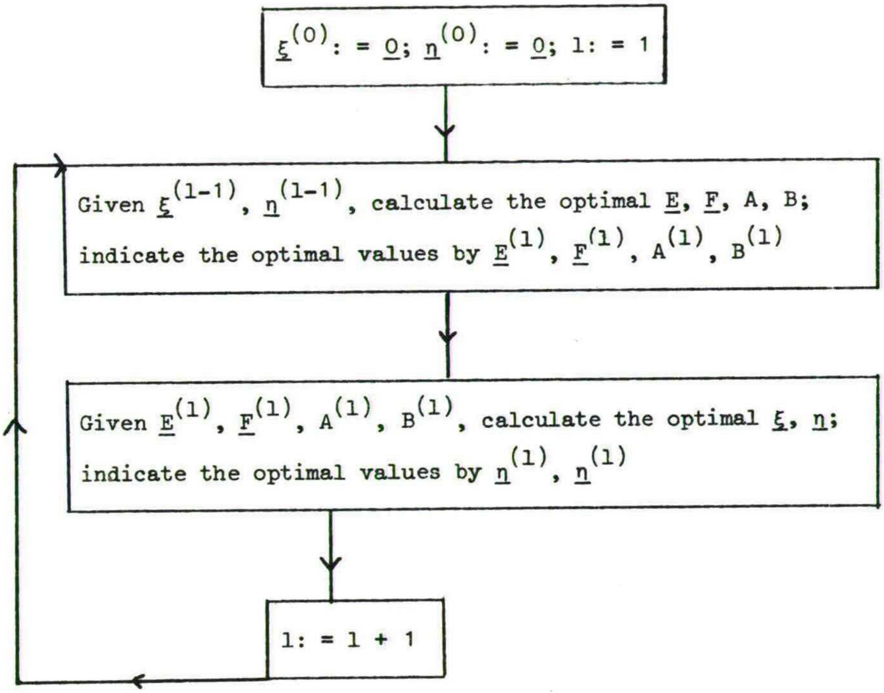
$$\sum_{t=0}^{T-1} \alpha_t^2(\tilde{y}_{t+1} - A\,\tilde{y}_t - B\,\tilde{x}_t)\tilde{x}_t' = 0 \qquad . \qquad (61)$$

From the equations (60) and (61) it is possible to calculate the optimal A and B; furthermore the optimal parameter fluctuations $\underline{E}$ and $\underline{F}$ are determined by the equations (58), (55) and (56).
It is easy to verify that the optimal matrices A and B are the solution of the following WLS problem:

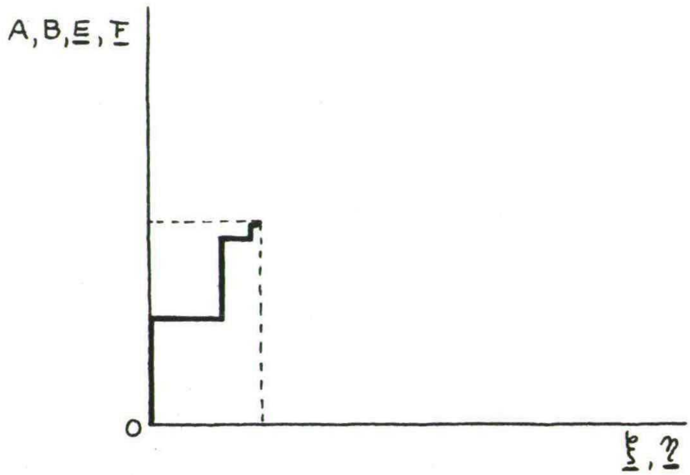
$$\min\{\sum_{t=0}^{T-1} \alpha_t^2\|\tilde{y}_{t+1} - A\,\tilde{y}_t - B\,\tilde{x}_t\|^2 \,|\, A \in M_{n,n},\ B \in M_{n,m}\}. \quad (62)$$

Hence the matrices A and B can be calculated using known technics from the linear regression theory. It is evident that some usual troubles may appear in the case that the solution of (62) is not unique. A necessary condition with respect to this uniqueness is given by: $T \geq n+m$.

So far the treatment of the two subproblems. We now have the possibility to solve the original problem in an iterative way. For that purpose we consider the following procedure:



This proces can be illustrated as follows:



It's evident that in principle any arbitrary point (A, B, E, F, ξ, η) can be taken as a starting point with respect to the iterative procedure.

3. A special case: all residuals equal to zero

In this chapter we shall investigate the problem that the unknown influences, described in chapter 1, are attributed fully to small observation errors in the data and to small time varying fluctuations on the structural matrices:

$$y_{t+1} + \xi_{t+1} = (A+E_t)(y_t+\xi_t) + (B+F_t)(x_t+\eta_t) \quad (t=0,\ldots,T-1). \tag{63}$$

Hence the residuals r_t do not appear here. Corresponding to the COLS problem we introduce here the following criterion:

$$\min \{ \sum_{t=0}^{T-1} [\|\xi_t\|^2 + \|\eta_t\|^2 + \|E_t\|^2 + \|F_t\|^2 + \|\xi_T\|^2 |$$

$$| y_{t+1} + \xi_{t+1} = (A+E_t)(y_t+\xi_t) + (B+F_t)(x_t+\eta_t),$$

$$A,E_t \in M_{n,n},\ B,F_t \in M_{n,m},\ \xi_t,\ \xi_T \in \mathbb{R}^n,$$

$$\eta_t \in \mathbb{R}^m \quad (t=0,\ldots,T-1)\}. \tag{64}$$

This minimization problem does not have the property that it can be stated as an unconstraint optimization problem by making a simple substitution. In the previous chapter we had that possibility. Hence we have to follow another method, namely the method of the Lagrange multipliers (see e.g. [6]). Define the following <u>Lagrange function</u>:

$$L(A,B,\xi_0,\ldots,\xi_T,\eta_0,\ldots,\eta_{T-1},E_0,\ldots,E_{T-1},F_0,\ldots,F_{T-1},\lambda_0,\ldots,\lambda_{T-1}) :=$$

$$\sum_{t=0}^{T-1} [\|\xi_t\|^2 + \|\eta_t\|^2 + \|E_t\|^2 + \|F_t\|^2] + \|\xi_T\|^2 +$$

$$+ 2 \sum_{t=0}^{T-1} \lambda_t'[y_{t+1}+\xi_{t+1}-(A+E_t)(y_t+\xi_t)-(B+F_t)(x_t+\eta_t)]. \tag{65}$$

Here the vectors $\lambda_t \in \mathbb{R}^n$ contain the nT Lagrange multipliers. The Lagrange theory says: derive the first order conditions of the function L with respect to A, B, $\xi_0,\ldots,\xi_T$, $\eta_0,\ldots,\eta_{T-1}$, $E_0,\ldots,E_{T-1}$, $F_0,\ldots,F_{T-1}$ and $\lambda_0,\ldots,\lambda_{T-1}$.
Then the set of extremal points of the problem (64) is a subset of the set of stationary points of the Lagrange problem (65).
A necessary condition for this assertion is the so-called <u>rank condition</u> i.e. the condition that the normals of the constraints in an extremal point of the problem (64) are linear independent. One can verify that in this case the rank condition holds.
First of all we shall calculate the gradients of the function L.
Again the shorter notation $L(A, B, \underline{\xi}, \underline{\eta}, \underline{E}, \underline{F}, \underline{\lambda})$ will be used.

$$\nabla_A L(A, B, \underline{\xi}, \underline{\eta}, \underline{E}, \underline{F}, \underline{\lambda}) = -2 \sum_{t=0}^{T-1} \lambda_t (y_t+\xi_t)' \tag{66}$$

$$\nabla_B L(A, B, \underline{\xi}, \underline{\eta}, \underline{E}, \underline{F}, \underline{\lambda}) = -2 \sum_{t=0}^{T-1} \lambda_t (x_t+\eta_t)' \tag{67}$$

$$\nabla_{\xi_0} L(A,B,\underline{\xi},\underline{\eta},\underline{E},\underline{F},\underline{\lambda}) = 2\,\xi_0 - 2(A+E_0)'\,\lambda_0 \tag{68}$$

$$\nabla_{\xi_t} L(A,B,\underline{\xi},\underline{\eta},\underline{E},\underline{F},\underline{\lambda}) = 2\,\xi_t - 2(A+E_t)'\,\lambda_t + 2\,\lambda_{t-1} \qquad (t=1,\ldots,T-1) \tag{69}$$

$$\nabla_{\xi_T} L(A,B,\underline{\xi},\underline{\eta},\underline{E},\underline{F},\underline{\lambda}) = 2\,\xi_T + 2\,\lambda_{T-1} \tag{70}$$

$$\nabla_{\eta_t} L(A,B,\underline{\xi},\underline{\eta},\underline{E},\underline{F},\underline{\lambda}) = 2\,\eta_t - 2(B+F_t)'\,\lambda_t \qquad (t=0,\ldots,T-1) \tag{71}$$

$$\nabla_{E_t} L(A,B,\underline{\xi},\underline{\eta},\underline{E},\underline{F},\underline{\lambda}) = 2\,E_t - 2\,\lambda_t (y_t+\xi_t)' \qquad (t=0,\ldots,T-1) \tag{72}$$

$$\nabla_{F_t} L(A,B,\underline{\xi},\underline{\eta},\underline{E},\underline{F},\underline{\lambda}) = 2\, F_t - 2\, \lambda_t(x_t+\eta_t)' \qquad (t=0,\ldots,T-1) \tag{73}$$

$$\nabla_{\lambda_t} L(A,B,\underline{\xi},\underline{\eta},\underline{E},\underline{F},\underline{\lambda}) =$$

$$= y_{t+1} + \xi_{t+1} - (A+E_t)(y_t+\xi_t) - (B+F_t)(x_t+\eta_t) \qquad (t=0,\ldots,T-1) \quad . \tag{74}$$

The first order conditions we find by setting all these gradients equal to zero:

$$\boxed{\sum_{t=0}^{T-1} \lambda_t(y_t+\xi_t)' = 0} \tag{75}$$

$$\boxed{\sum_{t=0}^{T-1} \lambda_t(x_t+\eta_t)' = 0} \tag{76}$$

$$\xi_0 = (A+E_0)'\, \lambda_0 \tag{77}$$

$$\xi_t = (A+E_t)'\, \lambda_t-\lambda_{t-1} \qquad (t=1,\ldots,T-1) \tag{78}$$

$$\xi_T = -\, \lambda_{T-1} \tag{79}$$

$$\boxed{\eta_t = (B+F_t)'\, \lambda_t} \qquad (t=0,\ldots,T-1) \tag{80}$$

$$\boxed{E_t = \lambda_t(y_t+\xi_t)'} \qquad (t=0,\dots,T-1) \tag{81}$$

$$\boxed{F_t = \lambda_t(x_t+\eta_t)'} \qquad (t=0,\dots,T-1) \tag{82}$$

$$\boxed{y_{t+1}+\xi_{t+1}=(A+E_t)(y_t+\xi_t)+(B+F_t)(x_t+\eta_t)} \quad . \tag{83}$$

$$(t=0,\dots,T-1)$$

The latter equations (83) are equal to the constraints of the original problem.
We notice an agreement of the conditions (75) ..... (82) with the conditions (28) ..... (35) in chapter 2. If we replace in (28) ..... (35) r_t by λ_t and put in (9) r_t equal to zero, we find the conditions (75) ...... (83)!

Remark. It follows again that the following equalities hold:

$$\sum_{t=0}^{T-1} E_t = 0 \tag{84}$$

and

$$\sum_{t=0}^{T-1} F_t = 0 \tag{85}$$

In the same way as in the previous chapter we consider here two subproblems, namely the problem of minimizing (64) with respect to $\underline{\xi}$ ánd $\underline{\eta}$ on one side, and with respect to A, B, $\underline{E}$ and $\underline{F}$ on the other side.

I. Fixed A, B, $\underline{E}$ and $\underline{F}$

Only the conditions (77), (78), (79), (80) and (83) are of importance. Because of the correspondence of this subproblem to the related subproblem of the previous chapter, we shall restrict ourselves to an enumeration of the results.
It is easy to verify that the optimal $\underline{\xi}$ and $\underline{\eta}$ can be calculated from the following equations:

$$\begin{bmatrix} I_n+A_0A_0'+B_0B_0' & -A_1' & & & 0 \\ -A_1 & I_n+A_1A_1'+B_1B_1' & -A_2' & & \\ & \ddots & \ddots & \ddots & \\ & -A_{T-2} & I_n+A_{T-2}'+B_{T-2}B_{T-2}' & -A_{T-1}' & \\ 0 & & -A_{T-1} & I_n+A_{T-1}A_{T-1}'+B_{T-1}B_{T-1}' \end{bmatrix} \begin{bmatrix} \lambda_0 \\ \lambda_1 \\ \vdots \\ \lambda_{T-1} \end{bmatrix} = \begin{bmatrix} s_0 \\ s_1 \\ \vdots \\ s_{T-1} \end{bmatrix} \quad (86)$$

$$\begin{cases} \xi_0 = A_0' \lambda_0 & (87) \\ \xi_t = A_t' \lambda_t - \lambda_{t-1} \quad (t=1,\dots,T-1) & (88) \\ \xi_T = -\lambda_{T-1} & (89) \end{cases}$$

$$\eta_t = B_t' \lambda_t \qquad (t=0,\dots,T-1) \quad . \qquad (90)$$

Furthermore A_t, B_t and s_t are again defined as follows:

$$A_t := A + E_t \qquad (t=0,\dots,T-1) \qquad (91)$$

$$B_t := B + F_t \qquad (t=0,\dots,T-1) \qquad (92)$$

$$s_t := y_{t+1} - A_t y_t - B_t x_t \quad (t=0,\dots,T-1) \quad . \qquad (93)$$

In Appendix B we shall prove that the block tridiagonal matrix in (86) is positive definite (and thus non-singular).

II. Fixed ξ and η

The conditions (75), (76), (81), (82) and (83) play a role here. Again we shall restrict ourselves to an enumeration of the results. The optimal structural matrices A and B can be calculated from the following WLS problem:

$$\min\{\sum_{t=0}^{T-1} \beta_t^2 \|\tilde{y}_{t+1} - A\,\tilde{y}_t - B\,\tilde{x}_t\|^2 \,|\, A \in M_{n,n},\ B \in M_{n,m}\}, \qquad (94)$$

where

$$\beta_t := [\,\|\tilde{y}_t\|^2 + \|\tilde{x}_t\|^2]^{-1/2} \qquad (t=0,\dots,T-1), \qquad (95)$$

$$\tilde{y}_t := y_t + \xi_t \qquad (t=0,\dots,T) \qquad (96)$$

and

$$\tilde{x}_t := x_t + \eta_t \qquad (t=0,\dots,T-1). \qquad (97)$$

Furthermore the optimal Lagrange multipliers are given by

$$\lambda_t = \beta_t^2(\tilde{y}_{t+1} - A\,\tilde{y}_t - B\,\tilde{x}_t) \quad (t=0,\dots,T-1), \qquad (98)$$

and the optimal parameter fluctuations are equal to

$$E_t = \lambda_t\,\tilde{y}_t' \qquad (t=0,\dots,T-1), \qquad (99)$$

$$F_t = \lambda_t\,\tilde{x}_t' \qquad (t=0,\dots,T-1). \qquad (100)$$

The treatment of the two subproblems implies again an iterative method for solving the original problem according to an analogue procedure as in chapter 2.

4. The general econometric model

Most of the current econometric models can not be written in the form defined by (1), since usually

(i) some exogenous variables are lagged one period

(ii) the structural equations are not stated in a reduced form

(iii) some elements of the structural matrices are constants (e.g. 0, 1 or -1).

In this chapter we shall extend the COLS method to models with the mentioned properties.
Hence we consider econometric models which have the following structural form:

$$A\, y_{t+1} + B\, y_t + C\, x_{t+1} + D\, x_t = 0 \quad (t=0,1,\dots\dots). \qquad (101)$$

Here the vectors $y_t \in \mathbb{R}^n$ and $x_t \in \mathbb{R}^m$ contain again respectively the endogenous and exogenous variables in time t. The matrices $A, B \in M_{n,n}$ and $C, D \in M_{n,m}$ are the structural matrices of the system.
We shall assume that the diagonal elements of the matrix A are all equal to one ($A_{ii} = 1$, $i = 1,\dots,n$). This normalization guarantees the uniqueness of (101).
Furthermore we assume that the whole set of unknown parameters appears in the first $n^{(1)}$ rows of the four matrices. The last $n^{(2)}$ equations ($n^{(1)} + n^{(2)} = n$) are identities (all parameters are known).
Hence in each of the first $n^{(1)}$ equations of (101) at least one unknown parameter appears.

Remark. The case that the matrix A is equal to the identity matrix, the matrix C is equal to the null matrix and $n^{(2)}$ is equal to zero, is corresponding to the model of the previous chapters.

The following partition of the matrices A, B, C and D is obvious:

$$A = \begin{bmatrix} A^{(1)} \\ \hdashline A^{(2)} \end{bmatrix}, \quad B = \begin{bmatrix} B^{(1)} \\ \hdashline B^{(2)} \end{bmatrix}, \quad C = \begin{bmatrix} C^{(1)} \\ \hdashline C^{(2)} \end{bmatrix}, \quad D = \begin{bmatrix} D^{(1)} \\ \hdashline D^{(2)} \end{bmatrix}, \qquad (102)$$

where $A^{(1)}, B^{(1)} \in M_{n^{(1)},n}$, $A^{(2)}, B^{(2)} \in M_{n^{(2)},n}$,

$$C^{(1)}, D^{(1)} \in M_{n^{(1)},m}, \; C^{(2)}, D^{(2)} \in M_{n^{(2)},m}.$$

Now it is possible to split (101) into two subsets of $n^{(1)}$ <u>reaction equations</u> and $n^{(2)}$ <u>definition equations</u>:

$$A^{(1)}y_{t+1} + B^{(1)}y_t + C^{(1)}x_{t+1} + D^{(1)}x_t = 0 \quad (t=0,1,\ldots.) \quad (103)$$

$$A^{(2)}y_{t+1} + B^{(2)}y_t + C^{(2)}x_{t+1} + D^{(2)}x_t = 0 \quad (t=0,1,\ldots.). \quad (104)$$

The matrices $A^{(1)}$, $B^{(1)}$, $C^{(1)}$ and $D^{(1)}$ contain the unknown parameters. All elements of the matrices $A^{(2)}$, $B^{(2)}$, $C^{(2)}$ and $D^{(2)}$ are known. The problem now is to determine the unknown parameters in (103) on the basis of observed values of the endogenous and exogenous variables during a time period $[0,\ldots,T]$. As in the previous chapters we introduce a residual r_t $(t=0,\ldots,T-1)$, observation errors ξ_t, η_t $(t=0,\ldots,T)$ in the data and time varying fluctuations E_t, F_t, G_t, H_t $(t=0,\ldots,T-1)$ on the structural matrices $A^{(1)}$, $B^{(1)}$, $C^{(1)}$, $D^{(1)}$ respectively:

$$(A^{(1)} + E_t)(y_{t+1} + \xi_{t+1}) + (B^{(1)} + F_t)(y_t + \xi_t) +$$

$$+ (C^{(1)} + G_t)(x_{t+1} + \eta_{t+1}) + (D^{(1)} + H_t)(x_t + \eta_t) +$$

$$+ r_t = 0 \qquad (t=0,1,\ldots,T-1) \qquad (105)$$

$$A^{(2)}(y_{t+1} + \xi_{t+1}) + B^{(2)}(y_t + \xi_t) + C^{(2)}(x_{t+1} + \eta_{t+1}) +$$

$$D^{(2)}(x_t + \eta_t) = 0 \quad (t=0,1,\ldots,T-1). \tag{106}$$

We notice that in the matrices E_t, F_t, G_t and H_t $(t=0,\ldots,T-1)$ now appear some elements equal to zero, namely those elements corresponding to the constant elements in respectively the matrices $A^{(1)}$, $B^{(1)}$, $C^{(1)}$ and $D^{(1)}$.
It follows that instead of (106) we can write

$$A^{(2)}\xi_{t+1} + B^{(2)}\xi_t + C^{(2)}\eta_{t+1} + D^{(2)}\eta_t = 0 \ (t=0,\ldots,T-1). \tag{107}$$

Analogous to the COLS method of chapter 2 we define here the following optimization problem for evaluating the unknown parameters of the model:

$$\min\{ \sum_{t=0}^{T-1} [\|r_t\|^2+\|\xi_t\|^2+\|\eta_t\|^2+\|E_t\|^2+\|F_t\|^2+\|G_t\|^2+\|H_t\|^2] +$$

$$+ \|\xi_T\|^2 + \|\eta_T\|^2 | (A^{(1)}+E_t)(y_{t+1}+\xi_{t+1}) + (B^{(1)}+F_t)(y_t+\xi_t) +$$

$$+ (C^{(1)}+G_t)(x_{t+1}+\eta_{t+1}) + (D^{(1)}+H_t)(x_t+\eta_t) + r_t = 0,$$

$$A^{(2)}\xi_{t+1} + B^{(2)}\xi_t + C^{(2)}\eta_{t+1} + D^{(2)}\eta_t = 0,\ E_t,\ F_t \in M_{n^{(1)},n},$$

$$G_t,\ H_t \in M_{n^{(1)},m},\ \xi_t \in \mathbb{R}^n,\ \eta_t \in \mathbb{R}^m,\ r_t \in \mathbb{R}^{n^{(1)}} \quad (t=0,\ldots,T-1),$$

$$\xi_T \in \mathbb{R}^n,\ \eta_T \in \mathbb{R}^m,\ A^{(1)},B^{(1)} \in M_{n^{(1)},n},\ C^{(1)},D^{(1)} \in M_{n^{(1)},m} \}. \tag{108}$$

Once again we notice that the optimization is not with respect to all elements of $A^{(1)}$, $B^{(1)}$, $C^{(1)}$ and $D^{(1)}$!

Similarly as in chapter 2 we consider two subproblems, which enables us to solve the original problem (108) in an iterative way.

I. Fixed $A^{(1)}$, $B^{(1)}$, $C^{(1)}$, $D^{(1)}$, $\underline{E}$, $\underline{F}$, $\underline{G}$ and $\underline{H}$

The minimization is only with respect to the errors in the data $\underline{\xi}$ and $\underline{\eta}$.
Define $A_t^{(1)} := A^{(1)} + E_t$, $B_t^{(1)} := B^{(1)} + F_t$, $C_t^{(1)} := C^{(1)} + G_t$, $D_t^{(1)} := D^{(1)} + H_t$ and $s_t := A_t^{(1)} y_{t+1} + B_t^{(1)} y_t + C_t^{(1)} x_{t+1} + D_t^{(1)} x_t$ $(t=0,\ldots,T-1)$.
Then the problem is:

$$\min\{ \sum_{t=0}^{T-1} \|r_t\|^2 + \sum_{t=0}^{T} [\|\xi_t\|^2 + \|\eta_t\|^2] \,|\, s_t + A_t^{(1)}\xi_{t+1} + B_t^{(1)}\xi_t +$$

$$+ C_t^{(1)}\eta_{t+1} + D_t^{(1)}\eta_t + r_t = 0,\ A^{(2)}\xi_{t+1} + B^{(2)}\xi_t +$$

$$+ C^{(2)}\eta_{t+1} + D^{(2)}\eta_t = 0,\ r_t \in \mathbb{R}^{n^{(1)}}\ (t=0,\ldots,T-1),$$

$$\eta_t \in \mathbb{R}^m, \xi_t \in \mathbb{R}^n\ (t=0,\ldots,T)\}. \tag{109}$$

It follows that (109) can be stated in the form:

$$\min\{ \sum_{t=0}^{T-1} \|A_t^{(1)}\xi_{t+1} + B_t^{(1)}\xi_t + C_t^{(1)}\eta_{t+1} + D_t^{(1)}\eta_t + s_t\|^2$$

$$+ \sum_{t=0}^{T} [\|\xi_t\|^2 + \|\eta_t\|^2] \,|\, A^{(2)}\xi_{t+1} + B^{(2)}\xi_t + C^{(2)}\eta_{t+1} +$$

$$+ D^{(2)}\eta_t = 0\ (t=0,\ldots,T-1),\ \xi_t \in \mathbb{R}^n,\ \eta_t \in \mathbb{R}^m\ (t=0,\ldots,T\},$$

$$(110)$$

by eliminating the residuals r_t $(t=0,\dots,T-1)$.
We shall follow again the method of the Lagrange multipliers for solving problem (110).
One can verify that the rank condition holds if we assume that one of the matrices $A^{(2)}A^{(2)'}$, $B^{(2)}B^{(2)'}$, $C^{(2)}C^{(2)'}$, $D^{(2)}D^{(2)'}$ is an element of PD.
It will appear that this assumption is again of importance later on.
Define the <u>Lagrange function</u>:

$$K(\xi_0,\dots,\xi_T,\ \eta_0,\dots,\eta_T,\ \lambda_0,\dots,\lambda_{T-1}) :=$$

$$\sum_{t=0}^{T}[\|\xi_t\|^2+\|\eta_t\|^2]+\sum_{t=0}^{T-1}\|A_t^{(1)}\xi_{t+1}+B_t^{(1)}\xi_t+C_t^{(1)}\eta_{t+1}+$$

$$+D_t^{(1)}\eta_t+s_t\|^2-2\sum_{t=0}^{T-1}\lambda_t'[A^{(2)}\xi_{t+1}+B^{(2)}\xi_t+C^{(2)}\eta_{t+1}+$$

$$+D^{(2)}\eta_t]. \qquad (111)$$

Here the vectors $\lambda_t \in \mathbb{R}^{n^{(2)}}$ contain the Lagrange multipliers $(t=0,\dots,T-1)$. We use the shorter notation $K(\underline{\xi},\ \underline{\eta},\ \underline{\lambda})$. The gradients of the function K are equal to

$$\left\|\begin{aligned}&\nabla_{\xi_0}K(\underline{\xi},\ \underline{\eta},\ \underline{\lambda}) = 2\,\xi_0 + \\ &\quad + 2\,B_0^{(1)'}[A_0^{(1)}\xi_1+B_0^{(1)}\xi_0+C_0^{(1)}\eta_1+D_0^{(1)}\eta_0+s_0] + \\ &\quad - 2\,B^{(2)'}\lambda_0\end{aligned}\right. \qquad (112)$$

$$\nabla_{\xi_t} K(\underline{\xi}, \underline{\eta}, \underline{\lambda}) = 2\,\xi_t +$$

$$+ 2\, B_t^{(1)'}[A_t^{(1)}\xi_{t+1} + B_t^{(1)}\xi_t + C_t^{(1)}\eta_{t+1} + D_t^{(1)}\eta_t + s_t] +$$

$$+ 2\, A_{t-1}^{(1)'}[A_{t-1}^{(1)}\xi_t + B_{t-1}^{(1)}\xi_{t-1} + C_{t-1}^{(1)}\eta_t + D_{t-1}^{(1)}\eta_{t-1} + s_{t-1}] +$$

$$-2\, A^{(2)'}\,\lambda_{t-1} - 2\, B^{(2)'}\,\lambda_t \quad (t=1,\ldots,T-1) \qquad (113)$$

$$\nabla_{\xi_T} K(\underline{\xi}, \underline{\eta}, \underline{\lambda}) = 2\,\xi_T +$$

$$+ 2\, A_{T-1}^{(1)'}[A_{T-1}^{(1)}\xi_T + B_{T-1}^{(1)}\xi_{T-1} + C_{T-1}^{(1)}\eta_T + D_{T-1}^{(1)}\eta_{T-1} + s_{T-1}] +$$

$$-\,2\, A^{(2)'}\,\lambda_{T-1} \qquad (114)$$

$$\nabla_{\eta_0} K(\underline{\xi}, \underline{\eta}, \underline{\lambda}) = 2\,\eta_0 +$$

$$+ 2\, D_0^{(1)'}[A_0^{(1)}\xi_1 + B_0^{(1)}\xi_0 + C_0^{(1)}\eta_1 + D_0^{(1)}\eta_0 + s_0] +$$

$$-\,2\, D^{(2)'}\,\lambda_0 \qquad (115)$$

$$\nabla_{\eta_t} K(\underline{\xi}, \underline{\eta}, \underline{\lambda}) = 2\,\eta_t +$$

$$+ 2\, D_t^{(1)'}[A_t^{(1)}\xi_{t+1} + B_t^{(1)}\xi_t + C_t^{(1)}\eta_{t+1} + D_t^{(1)}\eta_t + s_t] +$$

$$+ 2\, C_{t-1}^{(1)'}[A_{t-1}^{(1)}\xi_t + B_{t-1}^{(1)}\xi_{t-1} + C_{t-1}^{(1)}\eta_t + D_{t-1}^{(1)}\eta_{t-1} + s_{t-1}] +$$

$$-\,2\, D^{(2)'}\,\lambda_t - 2\, C^{(2)'}\,\lambda_{t-1} \quad (t=1,\ldots,T-1) \qquad (116)$$

$$\nabla_{\eta_T} K(\underline{\xi}, \underline{\eta}, \underline{\lambda}) = 2\ \eta_T +$$

$$+\ C_{T-1}^{(1)'}[A_{T-1}^{(1)}\ \xi_T + B_{T-1}^{(1)}\xi_{T-1} + C_{T-1}^{(1)}\eta_T + D_{T-1}^{(1)}\eta_{T-1} + s_{T-1}] +$$

$$-\ 2\ C^{(2)'}\ \lambda_{T-1} \tag{117}$$

$$\nabla_{\lambda_t} K(\underline{\xi}, \underline{\eta}, \underline{\lambda},) = A^{(2)}\xi_{t+1} + B^{(2)}\xi_t + C^{(2)}\eta_{t+1} + D^{(2)}\eta_t$$

$$(t=0,\dots,T-1). \tag{118}$$

Because $A_t^{(1)}\xi_{t+1} + B_t^{(1)}\xi_t + C_t^{(1)}\eta_{t+1} + D_t^{(1)}\eta_t + s_t = -r_t$ $(t=0,\dots,T-1)$ we can simplify these gradients as follows:

$$\nabla_{\xi_0} K(\underline{\xi}, \underline{\eta}, \underline{\lambda}) = 2\ \xi_0 - 2B_0^{(1)'}r_0 - 2\ B^{(2)'}\lambda_0 \tag{119}$$

$$\nabla_{\xi_t} K(\underline{\xi}, \underline{\eta}, \underline{\lambda}) = 2\ \xi_t - 2B_t^{(1)'}r_t - 2\ A_{t-1}^{(1)'}r_{t-1} +$$

$$-\ 2\ B^{(2)'}\lambda_t - 2\ A^{(2)'}\lambda_{t-1}\ (t=1,\dots,T-1) \tag{120}$$

$$\nabla_{\xi_T} K(\underline{\xi}, \underline{\eta}, \underline{\lambda}) = 2\ \xi_T - 2A_{T-1}^{(1)'}r_{T-1} - 2\ A^{(2)'}\ \lambda_{T-1} \tag{121}$$

$$\nabla_{\eta_0} K(\underline{\xi}, \underline{\eta}, \underline{\lambda}) = 2\ \eta_0 - 2\ D_0^{(1)'}r_0 - 2\ D^{(2)'}\lambda_0 \tag{122}$$

$$\nabla_{\eta_t} K(\underline{\xi}, \underline{\eta}, \underline{\lambda}) = 2\ \eta_t - 2\ D_t^{(1)'}r_t - 2\ C_{t-1}^{(1)'}r_{t-1} +$$

$$-\ 2\ D^{(2)'}\lambda_t - 2\ C^{(2)'}\lambda_{t-1}\ \ (t=1,\dots,T-1) \tag{123}$$

$$\nabla_{\eta_T} K(\underline{\xi}, \underline{\eta}, \underline{\lambda}) = 2\ \eta_T - 2\ C_{T-1}^{(1)'}r_{T-1} - 2\ C^{(2)'}\lambda_{T-1} \tag{124}$$

$$\nabla_{\lambda_t} K(\underline{\xi}, \underline{\eta}, \underline{\lambda}) = A^{(2)}\xi_{t+1} + B^{(2)}\xi_t + C^{(2)}\eta_{t+1} + D^{(2)}\eta_t \quad (t=0,\dots,T-1) \quad . \quad (125)$$

Setting these gradients equal to zero we get the stationarity conditions:

$$\xi_0 = B_0^{(1)'} r_0 + B^{(2)'}\lambda_0 \quad (126)$$

$$\xi_t = B_t^{(1)'} r_t + A_{t-1}^{(1)'} r_{t-1} + B^{(2)'}\lambda_t + A^{(2)'}\lambda_{t-1} \quad (t=1,\dots,T-1) \quad (127)$$

$$\xi_T = A_{T-1}^{(1)'} r_{T-1} + A^{(2)'}\lambda_{T-1} \quad (128)$$

$$\eta_0 = D_0^{(1)'} r_0 + D^{(2)'}\lambda_0 \quad (129)$$

$$\eta_t = D_t^{(1)'} r_t + C_{t-1}^{(1)'} r_{t-1} + D^{(2)'}\lambda_t + C^{(2)'}\lambda_{t-1} \quad (t=1,\dots,T-1) \quad (130)$$

$$\eta_T = C_{T-1}^{(1)'} r_{T-1} + C^{(2)'}\lambda_{T-1} \quad (131)$$

$$\boxed{A^{(2)}\xi_{t+1} + B^{(2)}\xi_t + C^{(2)}\eta_{t+1} + D^{(2)}\eta_t = 0} \quad . \quad (132)$$

Furthermore the relation

$$s_t + A_t^{(1)}\xi_{t+1} + B_t^{(1)}\xi_t + C_t^{(1)}\eta_{t+1} + D_t^{(1)}\eta_t + r_t = 0 \quad (133)$$

holds for $t=0,\dots,T-1$.

If we substitute the equations (126) ... (131) into the equations (133) we obtain:

$$\left\| \begin{aligned} & s_0 + A_0^{(1)} B_1^{(1)'} r_1 + A_0^{(1)} A_0^{(1)'} r_0 + A_0^{(1)} B^{(2)'} \lambda_1 + A_0^{(1)} A^{(2)'} \lambda_0 + \\ & + B_0^{(1)} B_0^{(1)'} r_0 + B_0^{(1)} B^{(2)'} \lambda_0 + \\ & + C_0^{(1)} D_1^{(1)'} r_1 + C_0^{(1)} C_0^{(1)'} r_0 + C_0^{(1)} D^{(2)'} \lambda_1 + C_0^{(1)} C^{(2)'} \lambda_0 + \\ & + D_0^{(1)} D_0^{(1)'} r_0 + D_0^{(1)} D^{(2)'} \lambda_0 + r_0 = 0, \end{aligned} \right. \tag{134}$$

$$\left\| \begin{aligned} & s_t + A_t^{(1)} B_{t+1}^{(1)'} r_{t+1} + A_t^{(1)} A_t^{(1)'} r_t + A_t^{(1)} B^{(2)'} \lambda_{t+1} + \\ & + A_t^{(1)} A^{(2)'} \lambda_t + B_t^{(1)} B_t^{(1)'} r_t + B_t^{(1)} A_{t-1}^{(1)} r_{t-1} + \\ & + B_t^{(1)} B^{(2)'} \lambda_t + B_t^{(1)} A^{(2)'} \lambda_{t-1} + C_t^{(1)} D_{t+1}^{(1)'} r_{t+1} + \\ & + C_t^{(1)} C_t^{(1)'} r_t + C_t^{(1)} D^{(2)'} \lambda_{t+1} + C_t^{(1)} C^{(2)'} \lambda_t + \\ & + D_t^{(1)} D_t^{(1)'} r_t + D_t^{(1)} C_{t-1}^{(1)'} r_{t-1} + D_t^{(1)} D^{(2)'} \lambda_t + \\ & + D_t^{(1)} C^{(2)'} \lambda_{t-1} + r_t = 0, \quad (t=1,\dots,T-2) \end{aligned} \right. \tag{135}$$

$$\left\| \begin{aligned} & s_{T-1} + A_{T-1}^{(1)} A_{T-1}^{(1)'} r_{T-1} + A_{T-1}^{(1)} A^{(2)'} \lambda_{T-1} + \\ & \quad + B_{T-1}^{(1)} B_{T-1}^{(1)'} r_{T-1} + B_{T-1}^{(1)} A_{T-2}^{(1)'} r_{T-2} + B_{T-1}^{(1)} B^{(2)'} \lambda_{T-1} \\ & \qquad\qquad + B_{T-1}^{(1)} A^{(2)'} \lambda_{T-2} + \\ & \quad + C_{T-1}^{(1)} C_{T-1}^{(1)'} r_{T-1} + C_{T-1}^{(1)} C^{(2)'} \lambda_{T-1} + \end{aligned} \right.$$

$$+ D_{T-1}^{(1)} D_{T-1}^{(1)'} r_{T-1} + D_{T-1}^{(1)} C_{T-2}^{(1)'} r_{T-2} + D_{T-1}^{(1)} D^{(2)'} \lambda_{T-1} + D_{T-1}^{(1)} C^{(2)'} \lambda_{T-2} + r_{T-1} = 0. \qquad (136)$$

Similarly, substituting the equations (126) ... (131) into (132), we obtain:

$$\begin{aligned} & A^{(2)} B_1^{(1)'} r_t + A^{(2)} A_0^{(1)'} r_0 + A^{(2)} B^{(2)'} \lambda_1 + A^{(2)} A^{(2)'} \lambda_0 + \\ & + B^{(2)} B_0^{(1)'} r_0 + B^{(2)} B^{(2)'} \lambda_0 + \\ & + C^{(2)} D_1^{(1)'} r_1 + C^{(2)} C_0^{(1)'} r_0 + C^{(2)} D^{(2)'} \lambda_1 + C^{(2)} C^{(2)'} \lambda_0 + \\ & + D^{(2)} D_0^{(1)'} r_0 + D^{(2)} D^{(2)'} \lambda_0 = 0, \end{aligned} \qquad (137)$$

$$\begin{aligned} & A^{(2)} B_{t+1}^{(1)'} r_{t+1} + A^{(2)} A_t^{(1)'} r_t + A^{(2)} B^{(2)'} \lambda_{t+1} + A^{(2)} A^{(2)'} \lambda_t + \\ & + B^{(2)} B_t^{(1)'} r_t + B^{(2)} A_{t-1}^{(1)'} r_{t-1} + B^{(2)} B^{(2)'} \lambda_t + B^{(2)} A^{(2)'} \lambda_{t-1} + \\ & + C^{(2)} D_{t+1}^{(1)'} r_{t+1} + C^{(2)} C_t^{(1)'} r_t + C^{(2)} D^{(2)'} \lambda_{t+1} + \\ & \qquad + C^{(2)} C^{(2)'} \lambda_t + \\ & + D^{(2)} D_t^{(1)'} r_t + D^{(2)} C_{t-1}^{(1)'} r_{t-1} + D^{(2)} D^{(2)'} \lambda_t + \\ & \qquad + D^{(2)} C^{(2)'} \lambda_{t-1} = 0, \quad (t=1,\dots,T-2) \end{aligned} \qquad (138)$$

$$\begin{aligned}
&A^{(2)}A_{T-1}^{(1)'}r_{T-1} + A^{(2)}A^{(2)'}\lambda_{T-1} + \\
&+ B^{(2)}B_{T-1}^{(1)'}r_{T-1} + B^{(2)}A_{T-2}^{(1)'}r_{T-2} + B^{(2)}B^{(2)'}\lambda_{T-1} + B^{(2)}A^{(2)'}\lambda_{T-2} + \\
&+ C^{(2)}C_{T-1}^{(1)'}r_{T-1} + C^{(2)}C^{(2)'}\lambda_{T-1} + \\
&+ D^{(2)}D_{T-1}^{(1)'}r_{T-1} + D^{(2)}C_{T-2}^{(1)'}r_{T-2} + D^{(2)}D^{(2)'}\lambda_{T-1} + \\
&\qquad\qquad + D^{(2)}C^{(2)'}\lambda_{T-2} = 0. \quad (139)
\end{aligned}$$

The equations (134), (135) and (136) can be written as:

$$\begin{aligned}
&(I_{n^{(1)}} + A_0^{(1)}A_0^{(1)'} + B_0^{(1)}B_0^{(1)'} + C_0^{(1)}C_0^{(1)'} + D_0^{(1)}D_0^{(1)'})r_0 + \\
&\qquad + (A_0^{(1)}B_1^{(1)'} + C_0^{(1)}D_1^{(1)'})r_1 + \\
&\qquad + (A_0^{(1)}A^{(2)'} + B_0^{(1)}B^{(2)'} + C_0^{(1)}C^{(2)'} + D_0^{(1)}D^{(2)'})\lambda_0 + \\
&\qquad + (A_0^{(1)}B^{(2)'} + C_0^{(1)}D^{(2)'})\lambda_1 = -s_0, \qquad (140)
\end{aligned}$$

$$\begin{aligned}
&(B_t^{(1)}A_{t-1}^{(1)'} + D_t^{(1)}C_{t-1}^{(1)'})r_{t-1} + \\
&+ (I_{n^{(1)}} + A_t^{(1)}A_t^{(1)'} + B_t^{(1)}B_t^{(1)'} + C_t^{(1)}C_t^{(1)'} + D_t^{(1)}D_t^{(1)'})r_t + \\
&+ (A_t^{(1)}B_{t+1}^{(1)'} + C_t^{(1)}D_{t+1}^{(1)'})r_{t+1} + \\
&+ (B_t^{(1)}A^{(2)'} + D_t^{(1)}C^{(2)'})\lambda_{t-1} + \\
&+ (A_t^{(1)}A^{(2)'} + B_t^{(1)}B^{(2)'} + C_t^{(1)}C^{(2)'} + D_t^{(1)}D^{(2)'})\lambda_t + \\
&+ (A_t^{(1)}B^{(2)'} + C_t^{(1)}D^{(2)'})\lambda_{t+1} = -s_t, \quad (t=1,\dots,T-2) \quad (141)
\end{aligned}$$

$$\left\|\begin{array}{l}
(B^{(1)}_{T-1}A^{(1)'}_{T-2} + D^{(1)}_{T-1}C^{(1)'}_{T-2})r_{T-2} + \\
+ (I_{n^{(1)}} + A^{(1)}_{T-1}A^{(1)'}_{T-1} + B^{(1)}_{T-1}B^{(1)'}_{T-1} + C^{(1)}_{T-1}C^{(1)'}_{T-1} + D^{(1)}_{T-1}D^{(1)'}_{T-1})r_{T-1} + \\
+ (B^{(1)}_{T-1}A^{(2)'} + D^{(1)}_{T-1}C^{(2)'})\lambda_{T-2} + \\
+ (A^{(1)}_{T-1}A^{(2)'} + B^{(1)}_{T-1}B^{(2)'} + C_{T-1}C^{(2)'} + D_{T-1}D^{(2)'})\lambda_{T-1} = -s_{T-1}.
\end{array}\right. \tag{142}$$

Analogously the equations (137), (138) and (139) can be written as:

$$\left\|\begin{array}{l}
(A^{(2)}A^{(1)'}_{0} + B^{(2)}B^{(1)'}_{0} + C^{(2)}C^{(1)'}_{0} + D^{(2)}D^{(1)'}_{0})r_0 + \\
+ (A^{(2)}B^{(1)'}_{1} + C^{(2)}D^{(1)'}_{1})r_1 + \\
+ (A^{(2)}A^{(2)'} + B^{(2)}B^{(2)'} + C^{(2)}C^{(2)'} + D^{(2)}D^{(2)'})\lambda_0 + \\
+ (A^{(2)}B^{(2)'} + C^{(2)}D^{(2)'})\lambda_1 = 0,
\end{array}\right. \tag{143}$$

$$\left\|\begin{array}{l}
(B^{(2)}A^{(1)'}_{t-1} + D^{(2)}C^{(1)'}_{t-1})r_{t-1} + \\
+ (A^{(2)}A^{(1)'}_{t} + B^{(2)}B^{(1)'}_{t} + C^{(2)}C^{(1)'}_{t} + D^{(2)}D^{(1)'}_{t})r_t + \\
+ (A^{(2)}B^{(1)'}_{t+1} + C^{(2)}D^{(1)'}_{t+1})r_{t+1} + \\
+ (B^{(2)}A^{(2)'} + D^{(2)}C^{(2)'})\lambda_{t-1} + \\
+ (A^{(2)}A^{(2)'} + B^{(2)}B^{(2)'} + C^{(2)}C^{(2)'} + D^{(2)}D^{(2)'})\lambda_t + \\
+ (A^{(2)}B^{(2)'} + C^{(2)}D^{(2)'})\lambda_{t+1} = 0, \quad (t=1,\dots,T-2)
\end{array}\right. \tag{144}$$

$$\left\| \begin{aligned} &(B^{(2)}A^{(1)'}_{T-2} + D^{(2)}C^{(1)'}_{T-2})r_{T-2} + \\ &+ (A^{(2)}A^{(1)'}_{T-1} + B^{(2)}B^{(1)'}_{T-1} + C^{(2)}C^{(1)'}_{T-1} + D^{(2)}D^{(1)'}_{T-1})r_{T-1} + \\ &+ (B^{(2)}A^{(2)'} + D^{(2)}C^{(2)'})\lambda_{T-2} + \\ &+ (A^{(2)}A^{(2)'} + B^{(2)}B^{(2)'} + C^{(2)}C^{(2)'} + D^{(2)}D^{(2)'})\lambda_{T-1} = 0. \end{aligned} \right. \tag{145}$$

The equations (140) ... (145) can be written as one matrix equation as follows:

$$\begin{bmatrix} R_{11} & R_{12} \\ R_{21} & R_{22} \end{bmatrix} \begin{bmatrix} \underline{r} \\ \underline{\lambda} \end{bmatrix} = \begin{bmatrix} -\underline{s} \\ \underline{0} \end{bmatrix} , \tag{146}$$

where

$$\underline{r}: = \begin{bmatrix} r_0 \\ r_1 \\ \vdots \\ r_{T-1} \end{bmatrix}, \ \underline{\lambda}: = \begin{bmatrix} \lambda_0 \\ \lambda_1 \\ \vdots \\ \lambda_{T-1} \end{bmatrix}, \ \underline{s}: = \begin{bmatrix} s_0 \\ s_1 \\ \vdots \\ s_{T-1} \end{bmatrix}.$$

Further the block tridiagonal matrices R_{ij} $(i,j = 1,2)$ are defined as:

$$R_{11} := \left[\begin{array}{ccc|cc}
I_{n(1)}+A_0^{(1)}A_0^{(1)'}+B_0B_0^{(1)'}+C_0C_0^{(1)'}+D_0D_0^{(1)'} & & & A_0^{(1)}B_1^{(1)'}+C_0^{(1)}D_1^{(1)'} & 0 \\
B_1^{(1)}A_0^{(1)'}+D_1^{(1)}C_0^{(1)'} & \ddots & & & \ddots \\
& \ddots & \ddots & & A_{T-2}^{(1)}B_{T-1}^{(1)'}+C_{T-2}D_{T-1}^{(1)'} \\
0 & & B_{T-1}^{(1)}A_{T-2}^{(1)'}+D_{T-1}^{(1)}C_{T-2}^{(1)'} & & A_{T-1}^{(1)}A_{T-1}^{(1)'}+B_{T-1}^{(1)}B_{T-1}^{(1)'}+C_{T-1}^{(1)}C_{T-1}^{(1)'}+D_{T-1}^{(1)}D_{T-1}^{(1)'}
\end{array}\right],$$

$$R_{12} := \left[\begin{array}{ccc|cc}
A_0^{(1)}A^{(2)'}+B_0^{(1)}B^{(2)'}+C_0^{(1)}C^{(2)'}+D_0^{(1)}D^{(2)'} & & & A_0^{(1)}B^{(2)'}+C_0^{(1)}D^{(2)'} & 0 \\
B_1^{(1)}A^{(2)'}+D_1^{(1)}C^{(2)'} & \ddots & & & \ddots \\
& \ddots & \ddots & & A_{T-2}^{(1)}B^{(2)'}+C_{T-2}^{(1)}D^{(2)'} \\
0 & & B_{T-1}^{(1)}A^{(2)'}+D_{T-1}^{(1)}C^{(2)'} & & A_{T-1}^{(1)}A^{(2)'}+B_{T-1}^{(1)}B^{(2)'}+C_{T-1}^{(1)}C^{(2)'}+D_{T-1}^{(1)}D^{(2)'}
\end{array}\right],$$

$R_{21} := R'_{12}$,

$$R_{22} := \left[\begin{array}{ccc|ccc}
A^{(2)}A^{(2)'}+B^{(2)}B^{(2)'}+C^{(2)}C^{(2)'}+D^{(2)}D^{(2)'} & & & A^{(2)}B^{(2)'}+C^{(2)}D^{(2)'} & & 0 \\
B^{(2)}A^{(2)'}+D^{(2)}C^{(2)'} & \ddots & & & \ddots & \\
& \ddots & \ddots & & & A^{(2)}B^{(2)}+C^{(2)}D^{(2)'} \\
0 & & B^{(2)}A^{(2)'}+D^{(2)}C^{(2)'} & A^{(2)}A^{(2)'}+B^{(2)}B^{(2)'}+C^{(2)}C^{(2)'}+D^{(2)}D^{(2)'} & &
\end{array}\right].$$

The matrix

$$\begin{bmatrix} R_{11} & R_{12} \\ R_{21} & R_{22} \end{bmatrix}$$

can be written as:

$$\left[\begin{array}{c|c} I_{n^{(1)}T} & 0 \\ \hline 0 & 0 \end{array}\right] + \left[\begin{array}{ccccc} B_0^{(1)} & A_0^{(1)} & & & 0 \\ & B_1^{(1)} & A_1^{(1)} & & \\ & & \ddots & \ddots & \\ 0 & & & B_{T-1}^{(1)} & A_{T-1}^{(1)} \\ \hline B^{(2)} & A^{(2)} & & & 0 \\ & B^{(2)} & A^{(2)} & & \\ & & \ddots & \ddots & \\ 0 & & & B^{(2)} & A^{(2)} \end{array}\right] \left[\begin{array}{ccccc} B_0^{(1)} & A_0^{(1)} & & & 0 \\ & B_1^{(1)} & A_1^{(1)} & & \\ & & \ddots & \ddots & \\ 0 & & & B_{T-1}^{(1)} & A_{T-1}^{(1)} \\ \hline B^{(2)} & A^{(2)} & & & 0 \\ & B^{(2)} & A^{(2)} & & \\ & & \ddots & \ddots & \\ 0 & & & B^{(2)} & A^{(2)} \end{array}\right]' +$$

$$
+ \left[\begin{array}{cccccc}
D_0^{(1)} & C_0^{(1)} & & & & 0 \\
& D_1^{(1)} & C_1^{(1)} & & & \\
& & \ddots & \ddots & & \\
0 & & & & D_{T-1}^{(1)} & C_{T-1}^{(1)} \\
\hline
D^{(2)} & C^{(2)} & & & & 0 \\
& D^{(2)} & C^{(2)} & & & \\
& & \ddots & \ddots & & \\
0 & & & & D^{(2)} & C^{(2)}
\end{array}\right]
\left[\begin{array}{cccccc}
D_0^{(1)} & C_0^{(1)} & & & & \\
& D_1^{(1)} & C^{(1)} & & & \\
& & \ddots & \ddots & & \\
& & & & D_{T-1}^{(1)} & C_{T-1}^{(1)} \\
\hline
D^{(2)} & C^{(2)} & & & & 0 \\
& D^{(2)} & C^{(2)} & & & \\
& & \ddots & \ddots & & \\
0 & & & & D^{(2)} & C^{(2)}
\end{array}\right]'.
$$

It is evident that the matrix

$$
\begin{bmatrix} R_{11} & R_{12} \\ R_{21} & R_{22} \end{bmatrix}
$$

is an element of PSD. In Appendix C we shall prove that this matrix is an element of PD assuming that one of the matrices

$$
A^{(2)}A^{(2)'}, \; B^{(2)}B^{(2)'}, \; C^{(2)}C^{(2)'}, \; D^{(2)}D^{(2)'}
$$

is an element of PD. In that case the matrix

$$
\begin{bmatrix} R_{11} & R_{12} \\ R_{21} & R_{22} \end{bmatrix}
$$

is non-singular and from the equations (146), (126) ... (131), (133) we can determine the optimal $\underline{\xi}$, $\underline{\eta}$ and $\underline{r}$.

II. Fixed $\underline{\xi}$ and $\underline{\eta}$.

The minimization is with respect to the parameters in $A^{(1)}$, $B^{(1)}$, $C^{(1)}$, $D^{(1)}$ and the parameter fluctuations $\underline{E}$, $\underline{F}$, $\underline{G}$, $\underline{H}$ only.
Define $\tilde{y}_t := y_t + \xi_t$ and $\tilde{x}_t := x_t + \eta_t$ $(t=0,\dots,T)$.

Then we have the following minimization problem:

$$\min\{\sum_{t=0}^{T-1} [\|r_t\|^2 + \|E_t\|^2 + \|F_t\|^2 + \|G_t\|^2 + \|H_t\|^2] \,|$$

$$|(A^{(1)}+E_t)\tilde{y}_{t+1} + (B^{(1)}+F_t)\tilde{y}_t + (C^{(1)}+G_t)\tilde{x}_{t+1} +$$

$$+ (D^{(1)}+H_t)\tilde{x}_t + r_t = 0,\ E_t, F_t \in M_{n^{(1)},n},$$

$$G_t, H_t \in M_{n^{(1)},m},\ r_t \in \mathbb{R}^{n^{(1)}} \quad (t=0,\dots,T-1),$$

$$A^{(1)}, B^{(1)} \in M_{n^{(1)},n},\ C^{(1)}, D^{(1)} \in M_{n^{(1)},m}\}. \tag{147}$$

Note that the equation (107) don't play a role here.
Furthermore some elements in the matrices $A^{(1)}$, $B^{(1)}$, $C^{(1)}$ and $D^{(1)}$ are known constants and the corresponding elements in the matrices E_t, F_t, G_t and H_t $(t=0,\dots,T-1)$ are zero.
The constraint

$$(A^{(1)}+E_t)\tilde{y}_{t+1} + (B^{(1)}+F_t)\tilde{y}_t + (C^{(1)}+G_t)\tilde{x}_{t+1} +$$

$$+ (D^{(1)}+H_t)\tilde{x}_t + r_t = 0 \qquad (t=0,\dots,T-1) \tag{148}$$

does consist of $n^{(1)}$ one-dimensional systems of equations. One can verify that from the independence of this $n^{(1)}$ systems of constraints, the minimization problem (147) can be split in $n^{(1)}$ onedimensional subproblems. Each of these subproblems has the following structure:

$$\min\{\sum_{t=0}^{T-1} [\rho_t^2 + \|e_t\|^2] \mid \rho_t = v_t-(a+e_t)'w_t,$$

$$\rho_t \in \mathbb{R},\ e_t \in \mathbb{R}^p\ (t=0,\ldots,T-1),\ a \in \mathbb{R}^p\}. \tag{149}$$

Here the constraint $\rho_t = v_t - (a+e_t)'w_t$ $(t=0,\ldots,T-1)$ corresponds to a certain single equation in (148). The vector $a \in \mathbb{R}^p$ contains all the unknown parameters in the corresponding rows of $A^{(1)}$, $B^{(1)}$, $C^{(1)}$ and $D^{(1)}$. The vector $e_t \in \mathbb{R}^p$ represents the related parameter fluctuation, while $\rho_t \in \mathbb{R}$ represents the corresponding component of the residual r_t $(t=0,\ldots,T-1)$. The minimization problem (149) can be written as:

$$\min\{\sum_{t=0}^{T-1} [(v_t-(a+e_t)'w_t)^2 + \|e_t\|^2] \mid a,e_0,\ldots,e_{T-1} \in \mathbb{R}^p\}. \tag{150}$$

Let us define the function

$$F(a,e_0,\ldots,e_{T-1}) := \sum_{t=0}^{T-1} [(v_t-(a+e_t)'w_t)^2 + \|e_t\|^2]. \tag{151}$$

We shall now investigate the problem of minimizing the function F with respect to $a, e_0,\ldots,e_{T-1} \in \mathbb{R}^p$.
We have the following gradients:

$$\nabla_a F(a,e_0,\ldots,e_{T-1}) = -2 \sum_{t=0}^{T-1} (v_t-(a+e_t)'w_t)w_t, \tag{152}$$

$$\nabla_{e_t} F(a,e_0,\dots,e_{T-1}) = 2\, e_t - 2(v_t-(a+e_t)'w_t)w_t. \qquad (153)$$

$$(t=0,\dots,T-1)$$

Hence the stationarity conditions are

$$\begin{cases} \sum_{t=0}^{T-1} \rho_t\, w_t = 0 & (154) \\ e_t = \rho_t\, w_t \quad (t=0,\dots,T-1). & (155) \end{cases}$$

Remark. The parameter fluctuations e_t satisfy the property: $\sum_{t=0}^{T-1} e_t = 0.$

We have

$$\rho_t = v_t-(a+e_t)'w_t = v_t-a'w_t - \|w_t\|^2 \rho_t,$$

thus we can write

$$\rho_t = \frac{v_t - a'w_t}{1+\|w_t\|^2} \quad (t=0,1,\dots,T-1). \qquad (156)$$

The condition (154) now is similar to

$$\sum_{t=0}^{T-1} \frac{(v_t-a'w_t)w_t}{1+\|w_t\|^2} = 0. \qquad (157)$$

It is easy to verify that (157) is equivalent to the stationarity condition for the following WLS problem:

$$\min\{ \sum_{t=0}^{T-1} \frac{(v_t-a'w_t)^2}{1+\|w_t\|^2} \,|\, a \in \mathbb{R}^p\}. \qquad (158)$$

Now usual technics from the linear regression theory can be applied determining the optimal parameter vector a. The optimal residuals and parameter fluctuations can be calculated from the equations (156) and (155) respectively.

Remark. Problem (158) is analogous to the WLS problem (62).

So far the treatment of the two subproblems. We notice again that we have now the possibility for solving the original problem iteratively.
In the next chapter we give same experimental results.

5. Some experimental results

In this chapter the theory of chapter 4 is applied to the Model I of L.R. Klein [3] .
This model is a system of 6 equation describing the American economy, three reaction equations and three identities. In the model six endogenous and three exogenous variables occur.
The unknown parameters appear in the reaction equaitons. The sample period is 1920-1941.
The concerning variables are:

C : consumption (endogenous)

Π : profits (endogenous)

W_1 : private wage bill (endogenous)

W_2 : government wage bill (exogenous)

I : net investment (endogenous)

K : end-of-year stock of capital (endogenous)

Y : net national income (endogenous)

T : business taxes (exogenous)

G : government expenditure plus net foreign balance (exogenous).

The six equations have the following form:

$$C = \alpha_0 + \alpha_1(W_1+W_2) + \alpha_2\Pi \tag{159}$$

$$I = \beta_0 + \beta_1\Pi + \beta_2\Pi_{-1} + \beta_3 K_{-1} \tag{160}$$

$$W_1 = \gamma_0 + \gamma_1(Y+T-W_2) + \gamma_2(Y+T-W_2)_{-1} + \gamma_3\, tm \tag{161}$$

$$Y + T = C + I + G \tag{162}$$

$$Y = W_1 + W_2 + \Pi \tag{163}$$

$$\Delta K = I \quad . \tag{164}$$

Here the index -1 means that the corresponding variable is lagged one year. Furthermore we have $\Delta K = K - K_{-1}$.
In the third equation the time tm can be considered to be an exogenous variable.
Following the argument of L.R. Klein we assign tm the values -10, -9,...,9,10. So 1921 corresponds to tm = -10, 1922 to tm = -9 etc.
The first three equations are the reaction equations, in which α_i, β_i and γ_i represent the unknown parameters. The latter equations are the identities. First of all we shall introduce time indices in the equations (159) ... (164).

Instead of the years 1921 - 1941 we shall consider from now on the periods 1 - 21.

$$C_{t+1} = \alpha_0 + \alpha_1[(W_1)_{t+1} + (W_2)_{t+1}] + \alpha_2\, \Pi_{t+1} \quad (t=0,\dots,20) \tag{165}$$

$$I_{t+1} = \beta_0 + \beta_1\, \Pi_{t+1} + \beta_2\, \Pi_t + \beta_3\, K_t \qquad (t=0,\dots,20) \tag{166}$$

$$\begin{aligned}(W_1)_{t+1} = \gamma_0 + \gamma_1[Y_{t+1} + T_{t+1} - (W_2)_{t+1}] &+ \\ + \gamma_2[Y_t + T_t - (W_2)_t] + \gamma_3(tm)_{t+1} &\qquad (t=0,\dots,20)\end{aligned} \tag{167}$$

$$Y_{t+1} + T_{t+1} = C_{t+1} + I_{t+1} + G_{t+1} \qquad (t=0,\dots,20) \qquad (168)$$

$$Y_{t+1} = (W_1)_{t+1} + (W_2)_{t+1} + \Pi_{t+1} \qquad (t=0,\dots,20) \qquad (169)$$

$$K_{t+1} = I_{t+1} + K_t \qquad (t=0,\dots,20) \quad . \qquad (170)$$

Next we shall write the equations (165) ... (170) in the form (103), (104). For that purpose we define the endogenous vector y_t and the exogenous vector x_t as follows:

$$y_t := \begin{bmatrix} C_t \\ I_t \\ (W_1)_t \\ Y_t \\ \Pi_t \\ K_t \end{bmatrix} \in \mathbb{R}^6 \qquad (t=0,\dots,21) \qquad (171)$$

and

$$x_t := \begin{bmatrix} G_t \\ (W_2)_t \\ T_t \\ (tm)_t \end{bmatrix} \in \mathbb{R}^4 \qquad (t=0,\dots,21) \qquad (172)$$

Then the equations (165) ... (170) can be written as:

$$- \begin{bmatrix} \alpha_0 \\ \beta_0 \\ \gamma_0 \end{bmatrix} + \begin{bmatrix} 1 & 0 & -\alpha_1 & 0 & -\alpha_2 & 0 \\ 0 & 1 & 0 & 0 & -\beta_1 & 0 \\ 0 & 0 & 1 & -\gamma_1 & 0 & 0 \end{bmatrix} y_{t+1} +$$

$$+ \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -\beta_2 & -\beta_3 \\ 0 & 0 & 0 & -\gamma_2 & 0 & 0 \end{bmatrix} y_t + \begin{bmatrix} 0 & -\alpha_1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & -\gamma_1 & -\gamma_1 & -\gamma_3 \end{bmatrix} x_{t+1} +$$

$$+ \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & \gamma_2 & -\gamma_2 & 0 \end{bmatrix} x_t = 0 \quad (t=0,\dots,20) \tag{173}$$

and

$$\begin{bmatrix} -1 & -1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & -1 & 1 & 0 \\ 0 & -1 & 0 & 0 & 0 & 1 \end{bmatrix} y_{t+1} + \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 \end{bmatrix} y_t +$$

$$\begin{bmatrix} -1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} x_{t+1} + \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} x_t = 0 \quad (t=0,\dots,20). \tag{174}$$

Remark. In equation (173) an inhomogenous term appears. The theory does not change essentially by the presence of this term!

The matrix

$$\begin{bmatrix} -1 & -1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & -1 & 1 & 0 \\ 0 & -1 & 0 & 0 & 0 & 1 \end{bmatrix}$$

corresponds to the matrix $A^{(2)}$ in the previous chapter. This matrix has rank 3, because the submatrix

$$\begin{bmatrix} -1 & -1 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{bmatrix}$$

is non-singular. Hence we conclude that the property

$$A^{(2)}A^{(2)'} \in PD$$

holds here and the theory of chapter 4 can be applied.
We notice that the number of unknown variables in the related minimization problem is equal to 525! (inclusive of the 63 residuals).

Before giving the results we shall make some remarks about the computer program leading to these results. The program is written in the language ALGOL 68. The least squares problem (158) is solved with standard algorithms form the linear regression theory based on Householder transformations.
The solution of the matrix equation (146) we computed with the conjugate gradient method of Fletcher-Reeves (see [7, p. 231]).
Full advantage is taken of the sparsity structure of the matrix

$$\begin{bmatrix} R_{11} & R_{21} \\ R_{21} & R_{22} \end{bmatrix}.$$

In Appendix D one can find the sample data during the years 1920 - 1941.

Estimation of the parameters with the OLS method gives the result:

$\alpha_0 = 16.430$	$\alpha_1 = 0.804$	$\alpha_2 = 0.251$	
$\beta_0 = 10.126$	$\beta_1 = 0.480$	$\beta_2 = 0.333$	$\beta_3 = -0.112$
$\gamma_0 = 1.497$	$\gamma_1 = 0.440$	$\gamma_2 = 0.146$	$\gamma_3 = 0.130$

- table 1 -

The corresponding OLS residuals are equal to:

	r_1	r_2	r_3
1921	-0.298	-0.067	-1.294
1922	-1.540	-0.048	0.296
1923	-1.573	1.247	1.188
1924	-0.423	-1.351	-0.136
1925	0.116	0.415	-0.465
1926	1.053	1.492	-0.484
1927	1.460	0.789	-0.728
1928	1.110	-0.632	0.339
1929	-0.469	1.083	1.196
1930	0.831	0.279	-0.151

- table 2^A -

	r_1	r_2	r_3
1931	0.033	0.037	0.594
1932	-0.147	0.366	0.103
1933	-0.138	0.224	0.450
1934	-0.223	-0.173	0.282
1935	-0.218	0.010	0.014
1936	1.342	0.972	-0.851
1937	-0.395	0.052	0.996
1938	0.352	-2.566	-0.469
1939	0.713	-0.687	-0.380
1940	0.694	-0.781	-1.091
1941	-2.279	-0.662	0.592

- table 2^B -

The sum of the squared residuals is equal to:

$$\sum_{t=0}^{20} \sum_{i=1}^{3} (r_t)_i^2 = 46.241 \qquad . \tag{175}$$

Estimation of the parameters with the iterative COLS method gives the result:

$\alpha_0 = 15.158$	$\alpha_1 = 0.842$	$\alpha_2 = 0.233$	
$\beta_0 = 10.455$	$\beta_1 = 0.481$	$\beta_2 = 0.333$	$\beta_3 = -0.114$
$\gamma_0 = 1.694$	$\gamma_1 = 0.438$	$\gamma_2 = 0.145$	$\gamma_3 = 0.141$

- table 3 -

The corresponding optimal parameter fluctuations are given in the tables 4^A and 4^B:

	$10^2\delta\alpha_0$	$10^2\delta\alpha_1$	$10^2\delta\alpha_2$	$10^2\delta\beta_0$	$10^2\delta\beta_1$	$10^2\delta\beta_2$	$10^2\delta\beta_3$
1921	0.012	0.338	0.148	0	-0.003	-0.004	-0.050
1922	-0.091	-2.917	-1.531	0	-0.004	-0.003	-0.043
1923	-0.081	-3.009	-1.496	0.004	0.065	0.059	0.648
1924	-0.013	-0.474	-0.249	-0.004	-0.073	-0.069	-0.709
1925	0.014	0.542	0.282	0.001	0.021	0.021	0.203
1926	0.055	2.220	1.069	0.004	0.073	0.075	0.737
1927	0.071	2.929	1.398	0.002	0.037	0.037	0.383
1928	0.049	2.087	1.026	-0.001	-0.030	-0.028	-0.294
1929	-0.022	-0.981	-0.470	0.002	0.053	0.051	0.510
1930	0.038	1.599	0.593	0.001	0.010	0.014	0.142
1931	0	0.004	0.001	0	0.002	0.002	0.033
1932	-0.005	-0.181	-0.037	0.001	0.006	0.010	0.186
1933	0.002	0.069	0.023	0.001	0.006	0.004	0.114
1934	-0.009	-0.330	-0.111	0	-0.005	-0.005	-0.082
1935	-0.012	-0.459	-0.164	0	0	0	0.005
1936	0.054	2.405	0.958	0.002	0.043	0.034	0.482
1937	-0.025	-1.191	-0.432	0	0.002	0.002	0.025
1938	0.006	0.269	0.090	-0.006	-0.095	-0.107	-1.252
1939	0.015	0.757	0.291	-0.002	-0.032	-0.026	-0.340
1940	0.010	0.504	0.201	-0.002	-0.040	-0.036	-0.382
1941	-0.068	-4.178	-1.589	-0.002	-0.036	-0.033	-0.316

- table 4^A -

	$10^2\delta\gamma_0$	$10^2\delta\gamma_1$	$10^2\delta\gamma_2$	$10^2\delta\gamma_3$
1921	-0.029	-1.337	-1.316	0.293
1922	0.008	0.383	0.348	-0.069
1923	0.022	1.230	1.077	-0.172
1924	-0.001	-0.058	-0.058	0.007
1925	-0.006	-0.348	-0.325	0.034
1926	-0.005	-0.341	-0.325	0.027
1927	-0.008	-0.521	-0.518	0.032
1928	0.005	0.302	0.301	-0.014
1929	0.014	0.959	0.923	-0.029
1930	-0.002	-0.093	-0.102	0.002
1931	0.009	0.474	0.543	0
1932	0.001	0.050	0.061	0.001
1933	0.010	0.427	0.419	0.019
1934	0.005	0.230	0.208	0.014
1935	-0.001	-0.055	-0.050	-0.004
1936	-0.013	-0.823	-0.714	-0.066
1937	0.012	0.748	0.722	0.069
1938	-0.007	-0.409	-0.436	-0.047
1939	-0.005	-0.362	-0.317	-0.042
1940	-0.011	-0.816	-0.749	-0.097
1941	0.004	0.361	0.309	0.041

- table 4^B -

In the tables 5^A and 5^B one can find the optimal disturbances on the endogenous and exogenous variables:

	$10^4\delta C$	$10^4\delta I$	$10^4\delta W_1$	$10^4\delta Y$	$10^4\delta\Pi$	$10^4\delta K$
1920	0	0	0	-0.369	-0.002	0.539
1921	-1.104	0.583	1.970	0.703	-1.578	1.122
1922	5.875	-1.155	-2.755	-1.586	3.372	-0.032
1923	5.299	-1.096	-3.502	-1.680	3.551	-1.126
1924	0.641	0.270	-0.257	-0.212	0.257	-0.859
1925	-0.907	-0.015	0.602	0.345	-0.624	-0.871
1926	-3.477	0.403	1.601	1.281	-1.966	-0.470
1927	-4.527	0.738	2.273	1.749	-2.712	0.267
1928	-2.963	0.627	1.089	1.109	-1.351	0.895
1929	1.736	-1.104	-1.482	-0.716	1.372	-0.210
1930	-2.420	0.611	1.137	0.926	-1.275	0.401
1931	0.143	-0.167	-0.507	-0.118	0.407	0.234
1932	0.446	-0.225	-0.176	-0.109	0.310	0.009
1933	-0.016	0.084	-0.448	-0.033	0.388	0.093
1934	0.625	-0.176	-0.498	-0.253	0.467	-0.083
1935	0.751	-0.451	-0.372	-0.298	0.312	-0.533
1936	-3.639	1.037	2.202	1.545	-2.265	0.503
1937	1.752	-0.730	-1.332	-0.749	1.099	-0.220
1938	-0.547	0.292	0.602	0.158	-0.753	0.065
1939	-1.101	0.253	0.657	0.345	-0.879	0.318
1940	-0.753	0.346	0.920	0.350	-0.857	0.665
1941	4.317	-1.281	-2.048	-1.256	2.113	-0.617

- table 5^A -

	$10^4 \delta G$	$10^4 \delta W_2$	$10^4 \delta T$	$10^4 \delta tm$
1920	0	0.369	-0.369	0
1921	0.040	0.224	-1.099	-0.405
1922	-2.926	-2.641	3.697	0.125
1923	-2.545	-2.144	3.635	0.342
1924	-0.622	-0.224	0.511	-0.011
1925	0.479	0.350	-0.773	-0.077
1926	1.985	1.453	-2.240	-0.057
1927	2.579	1.862	-2.741	-0.081
1928	1.896	1.188	-1.418	0.081
1929	-0.379	-0.665	1.017	0.205
1930	1.380	0.955	-1.285	-0.012
1931	0.154	-0.027	0.257	0.126
1932	-0.074	-0.254	0.266	0.016
1933	0.189	0.017	0.300	0.135
1934	-0.259	-0.229	0.488	0.065
1935	-0.410	-0.264	0.211	-0.013
1936	1.830	1.362	-2.149	-0.163
1937	-0.695	-0.587	1.128	0.168
1938	0.027	0.299	-0.377	-0.092
1939	0.408	0.530	-0.754	-0.067
1940	0.183	0.260	-0.549	-0.144
1941	-2.171	-1.764	2.396	0.072

- table 5^B -

The corresponding residuals are equal to:

	$10^4 r_1$	$10^4 r_2$	$10^4 r_3$
1921	1.144	-0.005	-2.806
1922	-8.801	-0.124	0.886
1923	-7.844	0.185	2.449
1924	-1.263	-0.437	-0.077
1925	1.386	0.102	-0.546
1926	5.462	0.357	-0.403
1927	7.105	0.163	-0.573
1928	4.859	-0.161	0.577
1929	-2.115	0.215	1.456
1930	3.800	0.055	-0.088

- table 6^A -

	$10^4 r_1$	$10^4 r_2$	$10^4 r_3$
1931	0.011	0.007	0.890
1932	-0.520	0.082	0.114
1933	0.205	0.050	0.954
1934	-0.884	-0.052	0.458
1935	-1.161	-0.012	-0.092
1936	5.469	0.232	-1.159
1937	-2.446	-0.026	1.179
1938	0.574	-0.625	-0.652
1939	1.510	-0.158	-0.473
1940	0.937	-0.179	-1.027
1941	-6.489	-0.274	0.509

- table 6^B -

The sum of the squared residuals, the squared disturbances on the data and the squared parameter fluctuations over all periods is equal to:

$$9.806_{10} - 3. \qquad (176)$$

Furthermore we find here for the sum of only the squared residuals:

$$3.762_{10} - 6 \quad . \qquad (177)$$

Note that the former number represents the minimal value of (108).

The above results were achieved in eight iterations.
One iteration corresponds to minimizing the objective function (108) with respect to the data disturbances on one side and with respect to the parameters and the parameter fluctuations on the other side.

After eight iterations the difference between two successive values of the objective function was less than $_{10}-8$. The total computation time turned out to be 694 seconds, i.e. $\pm$ 85 seconds for one iteration.

Further we notice that in the case the optimization is only with respect to the parameters $\alpha_0,\ldots,\gamma_3$ and the parameter fuluctuations (zero data disturbances), the minimal value of the objective function is very close to the value (176), namely

$$9.810_{10}-3. \tag{178}$$

The aim of achieving a significant reduction of the OLS residuals by introducing fluctuations on the data and the parameters appears to be realizable. With comparatively small fluctuations on the data and the parameters,the sum of squared residuals is reduced considerable.
Furthermore it appears that the corrections on the data have very little influence to this reduction. We conclude that for this sample data there exists, close to an autonomous model with parameters given in table 3, a time varying model with the properties:

(i) the residuals $(r_1)_t$, $(r_2)_t$, $(r_3)_t$ $(t=0,\ldots,20)$ lie in the interval $[-8.80_{10}-4,\ 7.11_{10}-4]$

(ii) the parameter fluctuations $(\delta\alpha_0)_t,\ldots,(\delta\gamma_3)_t$ $(t=0,\ldots,20)$ lie in the interval $[-4.18_{10}-2,\ 2.93_{10}-2]$.

6. The static model with stochastic parameter fluctuations

In this and the next chapter we discuss stochastic models. The identification method of the previous chapters is considered as a statistical estimation method with respect to the unknown parameters. We are interested in the statistical properties of the concerning estimators. Because the theory is rather complex, we restrict ourselves to models with stochastic residuals and stochastic parameter fluctuations only. Observation errors are left out of consideration.
Some literature about this subject one can find in [8], [9], [10], [11], [12, p.354] and [13, p. 622].
However, these references give only background information.
In this chapter we discuss static models. In chapter 7 some remarks are made concerning the dynamic model.

First of all consider the usual linear regression model:

$$\underline{y} = b'x + \underline{\varepsilon} \quad , \tag{179}$$

where $\underline{y}$ is a dependent, observable random variable, $x \in \mathbb{R}^p$ an observable, non-random vector of explanatory variables, the vector $b \in \mathbb{R}^p$ is a vector of unknown regression coefficients and $\underline{\varepsilon}$ a non-observable random error.
There are observations $y_1,\ldots,y_T$ and $x_1,\ldots,x_T$ available. Hence we can write:

$$\begin{aligned} \underline{y}_1 &= b'x_1 + \underline{\varepsilon} \\ &\vdots \\ \underline{y}_T &= b'x_T + \underline{\varepsilon}_T \end{aligned} \tag{180}$$

Suppose that

A1: $E\,\underline{\varepsilon}_t = 0 \quad (t=1,\dots,T)$ (181)

A2: $\mathrm{var}(\underline{\varepsilon}_t) = \sigma^2(1+x_t'Wx_t) \quad (t=1,\dots,T),$ (182)

where $W \in M_{p,p} \cap PD$ is a known matrix and σ^2 is an unknown parameter. The assumption A2 implies that we are dealing with observations in a heteroscedastic model.

Further we assume that $\underline{\varepsilon}_1,\dots,\underline{\varepsilon}_T$ are mutually independent:

A3: $\mathrm{cov}(\underline{\varepsilon}_t,\underline{\varepsilon}_s) = 0 \quad (t,s=1,\dots,T;\ t \neq s).$ (183)

From the assumptions A1 and A2 it follows that

$$E\,\underline{y}_t = b'x_t \qquad (t=1,\dots,T) \tag{184}$$

and

$$\mathrm{var}(\underline{y}_t) = \sigma^2(1+x_t'Wx_t) \qquad (t=1,\dots,T). \tag{185}$$

Now consider the following criterion generating an estimator $\hat{b}$ for the unknown parameter vector b:

$$\min\{G(b) \mid b \in \mathbb{R}^p\}, \tag{186}$$

where

$$G(b) := \sum_{t=1}^{T} \frac{(y_t - b'x_t)^2}{1+x_t'Wx_t} . \tag{187}$$

This method corresponds to a WLS method. From linear regression theory it is known that $\hat{b}$ is the best linear unbiased estimator (BLUE) for b. Furthermore

$$\hat{\sigma}^2 := \frac{1}{T-p} G(\hat{b}) \tag{188}$$

is an unbiased estimator for σ^2.

Remarks. 1°. The case W = 0 corresponds to the OLS method:

$$\min\{ \sum_{t=1}^{T} (y_t - b'x_t)^2 \,\big|\, b \in \mathbb{R}^p\}. \tag{189}$$

2°. A necessary and sufficient condition for the unique existence of the estimator $\hat{b}$ is the condition that the matrix

$$X := \begin{pmatrix} x_1' \\ \hline \vdots \\ \hline x_T' \end{pmatrix} \in M_{T,p}$$

has rank p. At any rate the condiiton $T \geq p$ must hold!

Nest consider the model

$$\underline{y} = (b+\underline{\delta})'x + \underline{e}, \tag{190}$$

where the vector $\underline{\delta} \in \mathbb{R}^p$ is a non-observable random fluctuation on the parameter vector b and $\underline{e}$ a non-observable random residual. Again we have observations $y_1,\ldots,y_T$ and $x_1,\ldots,x_T$:

$$\begin{aligned} \underline{y}_1 &= (b+\underline{\delta}_1)'x_1 + \underline{e}_1 \\ &\vdots \\ \underline{y}_T &= (b+\underline{\delta}_T)'x_T + \underline{e}_T \,. \end{aligned} \tag{191}$$

We shall analyse system (191) under the following set of assumptions:

A4: $\qquad E\,\underline{e}_t = 0 \qquad (t=1,\ldots,T)$ (192)

A5: $\qquad E\,\underline{\delta}_t = 0 \qquad (t=1,\ldots,T)$ (193)

A6: $\qquad \mathrm{var}(\underline{e}_t) = \sigma^2 \qquad (t=1,\ldots,T)$ (194)

A7: $\qquad \mathrm{VAR}(\underline{\delta}_t) = \sigma^2 W \qquad (t=1,\ldots,T)$ (195)

A8: $\qquad \mathrm{cov}(\underline{e}_t,\underline{\delta}_s) = 0 \qquad (t,s=1,\ldots,T)$ (196)

A9: $\qquad \mathrm{cov}(\underline{e}_t,\underline{e}_s) = 0 \qquad (t,s=1,\ldots,T;t\neq s)$ (197)

A10: $\qquad \mathrm{COV}(\underline{\delta}_t,\underline{\delta}_s) = 0 \qquad (t,s=1,\ldots,T;t\neq s).$ (198)

Here σ^2 is an unknown parameter and W is a known matrix, $W \in M_{p,p} \cap PD$.

From the assumptions A4 ... A8 it follows that

$$E\,\underline{y}_t = b'x_t \qquad (t=1,\ldots,T) \tag{199}$$

and

$$\mathrm{var}(\underline{y}_t) = \sigma^2(1+x_t'Wx_t) \quad (t=1,\ldots,T). \tag{200}$$

We now introduce the following criterion generating an estimator b^* for the vector b:

$$\min\{F(b,\delta_1,\ldots,\delta_T) | b,\delta_t \in \mathbb{R}^p \ (t=1,\ldots,T)\}, \tag{201}$$

where the function F is defined as

$$F(b,\delta_1,\ldots,\delta_T) := \sum_{t=1}^{T} [(y_t-(b+\delta_t)'x_t)^2 + \delta_t' W^{-1} \delta_t]. \tag{202}$$

Notice that the matrix W presents the possiblity of weighting the contribution of the parameter fluctuations δ_t with regard to the contribution of the residuals $y_t - (b+\delta_t)'x_t$ in criterion (201). Furthermore it is interesting that the matrix W is weighting the variance of $\underline{\delta}_t$ with respect to the variance of $\underline{e}_t$ in (194) and (195)!

It is clear that the criterion (201) is analogous to the criterion of chapter 2 (in the case the data disturbances are zero). Is it possible to make some judgements to the statistical properties of the estimator b^*?

The gradient of the function F with respect to the fluctuation δ_t $(t=1,\ldots,T)$ is equal to:

$$\nabla_{\delta_t} F(b,\delta_1,\ldots,\delta_T) = -2x_t(y_t-(b+\delta_t)'x_t) + 2\,W^{-1}\,\delta_t . \qquad (203)$$

Hence the stationarity condition can be written as

$$x_t(y_t-(b+\delta_t)'x_t) = W^{-1}\,\delta_t \quad (t=1,\ldots,T). \qquad (204)$$

Because $e_t = y_t - (b+\delta_t)'x_t$ it follows that

$$\delta_t = W\,x_t\,e_t \quad (t=1,\ldots,T) \qquad (205)$$

and thus

$$e_t = y_t - b'x_t - x_t'\,W\,x_t\,e_t \quad (t=1,\ldots,T). \qquad (206)$$

Hence we find:

$$e_t = \frac{y_t-b'x_t}{1+x_t'\,W\,x_t}$$

and from (205) we have:

$$\delta_t = \frac{Wx_t(y_t-b'x_t)}{1+x_t'Wx_t} =: \delta_t^*(b) \quad (t=1,\ldots,T). \tag{207}$$

Furthermore

$$F(b,\delta_1^*(b),\ldots,\delta_T^*(b)) =$$

$$= \sum_{t=1}^{T} \{[y_t-b'x_t - \frac{x_t'Wx_t(y_t-b'x_t)}{1+x_t'Wx_t}]^2 +$$

$$+ \frac{x_t'Wx_t(y_t-b'x_t)^2}{(1+x_t'Wx_t)^2}\} =$$

$$= \sum_{t=1}^{T} (y_t-b'x_t)^2\{(1-\frac{x_t'Wx_t}{1+x_t'Wx_t})^2 + \frac{x_t'Wx_t}{(1+x_t'Wx_t)^2}\} =$$

$$= \sum_{t=1}^{T} \frac{(y_t-b'x_t)^2}{1+x_t'Wx_t}. \tag{208}$$

Finally we find that

$$F(b_1\ \delta_1^*(b)_1,\ldots,\delta_T^*(b)) = G(b) \tag{209}$$

(see (187)) and we conclude that

$$b^* = \hat{b}! \tag{210}$$

Hence the criterions (186) and (201) generate the same estimator for b.

If we write

$$\underline{\delta}_t'x_t + \underline{e}_t =: \underline{\varepsilon}_t \qquad (t=1,\ldots,T) \tag{211}$$

then it follows that (191) is similar to (180) and the assumptions A1, A2 and A3 hold. Hence we find that b^* is an unbiased estimator for b. Further $\hat{\sigma}^2$ (see (188)) is an unbiased estimator for σ^2.

Above we made a remark that the criterion (186) corresponds to the OLS method in the case the matrix W is equal to the null matrix. Here we have the assumption that W is an element of PD.
It is not possible to substitute simply W=0 in the criterion (201). However we shall prove the OLS method can be considered as a limit case of (201). Namely, we shall indicate a sequence of matrices $(W_k)_{k \in \mathbb{N}}$, where $W_k \in PD(k \in \mathbb{N})$, $W_k \downarrow 0$ $(k \to \infty)$, with the property that the optimization problem

$$\min\{\sum_{t=1}^{T} [(y_t-(b+\delta_t)'x_t)^2 + \delta_t' W_k^{-1}\delta_t] \,|\, b,\delta_t \in \mathbb{R}^p \ (t=1,\ldots,T)\}$$

corresponds in the limit case to the OLS problem

$$\min\{\sum_{t=1}^{T} (y_t-b'x_t)^2 \,|\, b \in \mathbb{R}^p\}.$$

Define the following function

$$H(b) := \sum_{t=1}^{T} (y_t-b'x_t)^2. \tag{212}$$

Then the OLS method can be formulated as

$$\min\{H(b) \,|\, b \in \mathbb{R}^p\}. \tag{213}$$

Evidently the function H(b) can be written as a solution of an optimization problem with trivial constraints:

$$H(b) = \min\{\sum_{t=1}^{T} (y_t-(b+\delta_t)'x_t)^2 \,|\, \delta_t \in \mathbb{R}^p,\ \delta_t=0 \quad (t=1,\ldots.,T)\}. \tag{214}$$

It can be shown (see [6, p. 254]) that (214) can be written as

$$H(b) = \lim_{k\to\infty} C_k(b),$$

where the functions C_k ($k \in \mathbb{N}$) are defined as

$$C_k(b) := \min\{ \sum_{t=1}^{T} (y_t-(b+\delta_t)'x_t)^2 + \frac{1}{\alpha_k} \delta_t' W^{-1} \delta_t \mid \delta_t \in \mathbb{R}^p \ (t=1,\dots,T)\}. \quad (k \in \mathbb{N}) \tag{215}$$

Here the term $\delta_t' W^{-1} \delta_t$ represents a so-called <u>penalty function</u>, while the sequence $(\alpha_k)_{k \in \mathbb{N}}$ has the following properties:

(i) $\alpha_{k+1} < \alpha_k \qquad (k \in \mathbb{N})$

(ii) $\lim_{k\to\infty} \alpha_k = 0$

(iii) $\alpha_1 = 1$.

Hence the optimization problem (214) can be formulated as a limit case of a sequence of optimization problems <u>without</u> constraints. Analogue to the method in (208) one can find that

$$C_k(b) = \sum_{t=1}^{T} \frac{(y_t-b'x_t)^2}{1+\alpha_k x_t' W x_t}. \tag{216}$$

We notice that $C_1(b) = G(b)$ (see (187)).
Now one can easily verify the following property for these functions:

$$\min\{H(b) \mid b \in \mathbb{R}^p\} = \lim_{k\to\infty} \min\{C_k(b) \mid b \in \mathbb{R}^p\}. \tag{217}$$

If we define the sequence of matrices $(W_k)_{k \in \mathbb{N}}$ as

$$W_k := \alpha_k W \tag{218}$$

then we have $W_k \in PD$ and $W_k \to 0$ $(k \to \infty)$, while

$$\min\{H(b) \mid b \in \mathbb{R}^p\} =$$

$$= \lim_{k\to\infty} \min\{ \sum_{t=1}^{T} (y_t-(b+\delta_t)'x_t) + \delta_t' W_k^{-1} \delta_t \mid$$

$$b, \delta_t \in \mathbb{R}^p \quad (t=1,\ldots\ldots,T)\}. \tag{219}$$

This follows from (215) and (217).

7. The dynamic model with stochastic parameter fluctuations

In this chapter we discuss models of the form

$$\underline{y}_t = (a+\underline{\gamma}_t)\underline{y}_{t-1} + (b+\underline{\delta}_t)'x_t + \underline{e}_t. \quad (t=1,2,\ldots) \tag{220}$$

Here $\underline{y}_t$ is the endogenous, observable random variable at time t, $x_t \in \mathbb{R}^p$ the observable, non-random vector of exogenous variables at time t and $\underline{e}_t$ the non-observable random residual at time t. Further the variable a and the vector $b \in \mathbb{R}^p$ are the unknown structural parameters of the model. The random, non-observable variable $\underline{\gamma}_t$ and the random, non-observable vector $\underline{\delta}_t \in \mathbb{R}^p$ represent the parameter fluctuations at the time t on respectively a and b.

We shall analyse system (220) under the following set of assumptions:

A1: $$E\,\underline{e}_t = 0 \tag{221}$$

A2: $$E\,\underline{\gamma}_t = 0 \tag{222}$$

A3: $$E\,\underline{\delta}_t = 0 \tag{223}$$

A4: $$\mathrm{var}(\underline{e}_t) = \sigma^2 \tag{224}$$

A5: $$\mathrm{var}(\underline{\gamma}_t) = \sigma^2 v \tag{225}$$

A6: $$\mathrm{VAR}(\underline{\delta}_t) = \sigma^2 W \quad (t=1,2,\ldots) \tag{226}$$

A7: $$\mathrm{cov}(\underline{e}_t,\underline{e}_s) = 0 \quad (t,s=1,2,\ldots.;\ t\neq s) \tag{227}$$

A8: $$\mathrm{cov}(\underline{e}_t,\underline{\gamma}_s) = 0 \tag{228}$$

A9: $$\mathrm{cov}(\underline{e}_t,\underline{\delta}_s) = 0 \quad (t,s=1,2,\ldots) \tag{229}$$

A10: $\quad \mathrm{cov}(\underline{\gamma}_t, \underline{\gamma}_s) = 0 \quad (t,s=1,2,\ldots;t\neq s)$ (230)

A11: $\quad \mathrm{cov}(\underline{\gamma}_t, \underline{\delta}_s) = 0 \quad (t,s=1,2,\ldots)$ (231)

A12: $\quad \mathrm{COV}(\underline{\delta}_t, \underline{\delta}_s) = 0 \quad (t,s=1,2,\ldots;t\neq s)$. (232)

Here σ^2 is an <u>unknown</u> parameter. The matrix $W \in M_{p,p} \cap PD$ is known, just as the variable v ($v > 0$).
We notice that v and W give the possibility of weighting the variance of respectively $\underline{\gamma}_t$ and $\underline{\delta}_t$ with regard to the variance of $\underline{e}_t$.

The structural parameters a and b and the variance parameter σ^2 will be estimated from sample data.
The sample period is $[0,1,\ldots,T]$. Hence the observed values of $y_0, y_1, \ldots, y_T$ and $x_1, \ldots, x_T$ are given.

We introduce the following criterion:

$$\min\{ \sum_{t=1}^{T} [(y_t-(a+\gamma_t)y_{t-1}-(b+\delta_t)'x_t)^2 + v^{-1}\,\gamma_t^2 + \delta_t'\,W^{-1}\,\delta_t] \,|$$

$$|\, a,\gamma_t \in \mathbb{R}\,,\ b,\delta_t \in \mathbb{R}^p \ (t=1,\ldots,T)\}. \tag{233}$$

The minimization problem (233) generates estimators $\hat{a}$ and $\hat{b}$ for respectively a and b.
In an analogue manner as in the previous chapter one can verify that these estimators are the solution of the following WLS problem:

$$\min\{ \sum_{t=1}^{T} \frac{(y_t-ay_{t-1}-b'x_t)^2}{(1+vy_{t-1}^2+x_t'Wx_t} \,|\, a \in \mathbb{R}\,,\ b \in \mathbb{R}^p\}. \tag{234}$$

An obvious estimator for σ^2 is given by

$$\hat{\sigma}^2 := \frac{1}{T-p-1} \sum_{t=1}^{T} \frac{(y_t - ay_{t-1} - b'x_t)^2}{1 + vy_{t-1}^2 + x_t' W x_t} . \tag{239}$$

In models, with lagged endogenous variables, a least squares criterion like (234) generates in general inconsistent estimators. It is of importance to know if under certain conditions the estimators $\hat{a}$ and $\hat{b}$ have the asymptotical property of consistency after all. An exact answer to this question requires an extensive study of the theory of stochastic processes (stationarity properties and asymptotical properties of autocorrelation functions) and rather belongs beyond the framework of this research.
Therefore we restrict ourselves to an indication how the problem can be investigated.
In the case $v = 0$ and $W = 0$ the model (220) corresponds to

$$\underline{y}_t = a\underline{y}_{t-1} + b'x_t + \underline{e}_t \quad (t=1,2,\ldots,). \tag{240}$$

The assumptions A1, A4 and A7 hold.
Analogue to the theory at the end of the previous chapter the objective function (233) corresponds in the limit case to

$$\min\{\sum_{t=1}^{T} (y_t - ay_{t-1} - b'x_t)^2 \mid a \in \mathbb{R} , b \in \mathbb{R}^p\}. \tag{241}$$

This is in agreement with (234). Hence the estimators $\hat{a}$ and $\hat{b}$ are the OLS estimators for a and b and the estimator for σ^2 is given by

$$\hat{\sigma}^2 = \frac{1}{T-p-1} \sum_{t=1}^{T} (y_t - ay_{t-1} - b'x_t)^2. \tag{242}$$

In [14, p. 164] this model is discussed in detail, specially with regard to the asymptotical properties of the estimators (241), (242) for a, b and σ^2.
Under certain conditions these estimators are consistent.

These conditions exist of some stationarity conditions with respect to the stochastic process $\{\underline{y}_t\}_{t \geq 0}$.
Asymptotical properties of the autocorrelation function are translated in asymptotical properties of the estimators.
Further the following assumptions are made:

(i) the residuals $\underline{e}_t$ $(t=1,2,\ldots)$ are independently and identically distributed
(ii) the sequence $\{x_t\}_{t\in\mathbb{N}}$ is bounded
(iii) $|\hat{a}| \leq 1$.

If v = 0 and W ∈ PD, the model (220) can be written as

$$\underline{y}_t = a\underline{y}_{t-1} + (b+\underline{\delta}_t)'x_t + \underline{e}_t \quad (t=1,2,\ldots). \tag{243}$$

Here we have the following assumptions: A1, A3, A4, A6, A7, A9 and A12.
The objective function is given by

$$\min\{\sum_{t=1}^{T} [(y_t-ay_{t-1}-(b+\delta_t)'x_t)^2 + \delta_t' W^{-1} \delta_t] \,|\, a \in \mathbb{R},$$
$$b,\delta_t \in \mathbb{R}^p \quad (t=1,\ldots,T)\}. \tag{244}$$

We can write for (243):

$$\underline{y}_t = a\underline{y}_{t-1} + b'x_t + \underline{\varepsilon}_t \quad (t=1,2,\ldots,T), \tag{245}$$

where

$$E\,\underline{\varepsilon}_t = 0 \qquad (t=1,2,\ldots,T) \tag{246}$$

and

$$\mathrm{var}(\underline{\varepsilon}_t) = \sigma^2(1+x_t' \, W \, x_t) \quad (t=1,2,\ldots,T). \tag{247}$$

Further

$$\mathrm{cov}(\underline{\varepsilon}_t,\underline{\varepsilon}_s) = 0 \qquad (t,s=1,2,\ldots,T;t\neq s). \tag{248}$$

The estimators for a and b are again the solution of a WLS problem:

$$\min\{ \sum_{t=1}^{T} \frac{(y_t-ay_{t-1}-b'x_t)^2}{1+x_t'Wx_t} \,|\, a \in \mathbb{R}\,,\ b \in \mathbb{R}^{p}\}. \tag{249}$$

An estimator of σ^2 is given by

$$\hat{\sigma}^2 = \frac{1}{T-p-1} \sum_{t=1}^{T} \frac{(y_t-ay_{t-1}-b'x_t)}{1+x_t'Wx_t}. \tag{250}$$

We are dealing with a heteroscedastic model: the residuals $\underline{\varepsilon}_t$ are mutually independent, but they are not identically distributed. Anderson [14] requires the identical distribution of the residuals. Hence we conclude that the investigation of finding conditions with regard to the consistency of the estimators (249), (250), can be reduced to finding conditions which will allow to relax the assumption of identical distribution of the residuals.

We return now to the model (220) and the criterion (234). The model can be written as

$$\underline{y}_t = a\underline{y}_{t-1} + b'x_t + \underline{\varepsilon}_t \quad (t=1,2,\ldots.,T), \tag{251}$$

where the error term $\underline{\varepsilon}_t$ is given by

$$\underline{\varepsilon}_t := \underline{\gamma}_t \, \underline{y}_{t-1} + \underline{\delta}_t' \, x_t + \underline{e}_t \quad (t=1,2,\ldots,T). \tag{252}$$

The following difficulties arise, namely in (252) a product of

two stochastic variables appears and in (234) the term y_{t-1} appears in the denominator.
Probably the investigation of finding conditions for the estimators (234), (239) to satisfy the asymptotic property of consistency, is very complicated here.
Strong stationarity conditions will be necessary with regard to the stochastic process $\{\underline{y}_t\}_{t\geq 0}$.

8. Conclusions, remarks and some futher research

In the chapters 4 and 5 we discussed how the COLS method can be applied to general econometric models. The example (Model I of L.R. Klein) is a small model (6 endogenous variables, 4 exogenous variables and 21 periods). However, the number of unknown parameters in the minimization problem amounts to 525. One can imagine how this number increases when bigger models are considered. E.g. in the Klein-Goldberger Model 20 endogenous variables, 13 exogenous variables and 29 periods occur. This results in about 2500 unknown parameters!

The Model I of L.R. Klein has the property that a significant reduction of the OLS residuals is achieved by superposing comparatively small fluctuations on the data and the apremeters.

On account of the number of degrees of freedom (525 versus 74) it is evident that we can expect a reduction. But it is surprising that this reduction is so radical with relatively very small fluctuations.

We will not give an explanation for this property. However, we wonder if this property tells something about the validity of the model.

We conclude that for the data of Appendix D there exists close to an autonomous model a time varying model with an essential smaller residual part.

The COLS method (and also the CWLS method, which enables us to consider the errors relatively) is an attempt of charging parameter fluctuations and observation errors approximatively. We have not seen this in literature before.

A statistical description of models without observation errors turned out to be reasonable in the static case, i.e. concerning models without a time lagging structure. In the dynamic case a statistical description becomes very complicated.

In the structural form of the model (1) we did not specify the way the endogenous and exogenous variables are expressed: relatively or absolutely. We shall discuss these two cases with relation to the errors on the endogenous and exogenous variables on one side and to the homogeneity of the COLS problem on the other side.
First we consider the case the model (1) describes the relative growth of endogenous and exogenous variables. Then the COLS method (6) is a homogenous problem with respect to y_t and x_t. Namely, multiplication by a factor $k > 0$ of all the absolute sample data $\tilde{y}_0,\ldots,\tilde{y}_T$, $\tilde{x}_0,\ldots,\tilde{x}_{T-1}$ does not influence the relative sample data $y_0,\ldots,y_T$, $x_0,\ldots,x_{T-1}$ and hence this multiplication does not influence the objective function (6). However, observation errors are superposed on absolute variables and it is not clear what is the meaning of ξ_t and η_t in (5) when y_t and x_t are relative variables.

If z_t represents a relative growth of some variable ($t=0,\ldots,T$) then we have

$$z_t = \frac{\tilde{z}_t - \tilde{z}_{t-1}}{\tilde{z}_{t-1}} \quad (t=1,\ldots,T), \tag{253}$$

where $\tilde{z}_t$ is the notation for the related absolute variable. We shall prove that a given error ν_t on z_t implies a class of errors μ_t on $\tilde{z}_t$. The errors μ_t are corresponding to real observation errors. Suppose:

$$z_t + \nu_t = \frac{\tilde{z}_t+\mu_t-(\tilde{z}_{t-1}+\mu_{t-1})}{\tilde{z}_{t-1}+\mu_{t-1}} \quad (t=1,\ldots,T). \tag{254}$$

This equation can be written as

$$\mu_t - (1+z_t+\nu_t)\mu_{t-1} = \nu_t\,\tilde{z}_{t-1} \quad (t=1,\ldots,T). \tag{255}$$

This difference equation defines a class of errors μ_t on the

absolute variables $\tilde{z}_t$.
In this class we can choose an optimal element, e.g. according to the following criterion:

$$\min\{\sum_{t=0}^{T} \mu_t^2 | \mu_t - (1+z_t+\nu_t)\mu_{t-1} = \nu_t \tilde{z}_{t-1}, \mu_t, \mu_0 \in \mathbb{R} \quad (t=1,\dots,T)\}. \tag{256}$$

Next we discuss the case the equation (1) is dealing with <u>absolute</u> values of the data. Then the COLS problem (6) is not homogeneous. One can verify that the problem can be made homogenous by charging the residuals and the observation errors in a relative way:

$$\min\{\sum_{t=0}^{T-1} [\frac{\|r_t\|^2}{\|y_{t+1}\|^2} + \frac{\|\xi_t\|^2}{\|y_t\|^2} + \frac{\|\eta_t\|^2}{\|x_t\|^2} + \|E_t\|^2 + \|F_t\|^2] + \frac{\|\xi_T\|^2}{\|y_T\|^2} |$$

$$|y_{t+1} + \xi_{t+1} = (A+E_t)(y_t+\xi_t) + (B+F_t)(x_t+\eta_t) + r_t,$$

$$A, E_t \in M_{n,n}, \; B, F_t \in M_{n,m}, \; \xi_t, \xi_T, r_t \in \mathbb{R}^n,$$

$$\eta_t \in \mathbb{R}^m \quad (t=0,\dots,T-1)\}. \tag{257}$$

This criterion implies a CWLS problem!

Finally we make some suggestions for further research regarding the subject of this report:

1°. It is of importance to determine classes of problems for which the COLS method implies a considerable reduction of the OLS residuals.

2°. In chapter 2 we did not prove the existence of a minimum. Is it possible to construct a time series $\{y_0,\dots,y_T, x_0,\dots,x_{T-1}\}$

such that the COLS method has a solution for which some of the elements of A and B are equal to infinity?

3°. In some econometric models the endogenous vector y_t is not fully observable, but just some of its components or some linear combinations of its components. In the structural form (1) an <u>observable part</u> appears:

$$\begin{cases} y_{t+1} = Ay_t + Bx_t \\ z_t = Cy_t \qquad\qquad (t=0,1,\ldots.). \end{cases}$$

Can we extend the COLS theory to this kind of models?

4°. In the dynamic case the statistical properties of the estimators should be investigated (see chapter 7).

5°. The continuous analogy of the discrete model (1):

$$\dot{y}(t) = Ay(t) + Bx(t) \qquad (0 \leq t \leq T),$$

where $y:[0,T] \to \mathbb{R}^n$ and $x:[0,T] \to \mathbb{R}^m$ are functions, can be investigated.

References

[1] E. Malinvaud, "Méthodes statistiques de l'économétrie", Dunod, Paris, 1964.

[2] P. Schönfeld, "Methoden der oekonometrie", Verlag Vahlen, Berlin, 1969/71.

[3] L.R. Klein, "Economic fluctuations in de United States 1921-1941", Wiley, New York, 1950.

[4] W. Murray, "Numerical methods for unconstrained minimization problems", Academic Press, London, 1972.

[5] J. Kowalik/M.R. Osborne, "Methods for unconstrained optimization problems", American Elsevier Publishing Cy, New York, 1968.

[6] W.I. Zangwill, "Nonlinear programming, a unified approach", Prentice Hall, Englewood Cliffs, N.J., 1969.

[7] J.K. Reid, "Large sparse sets of linear equations", Academic Press, London, 1971.

[8] C. Rao, "The theory of least squares when the parameters are stochastic", Biometrica (1965), 52, p. 447.

[9] P.A.V.B. Swamy, "Efficient inference in a random coefficient regression model", Econometrica (1970), 38, p. 311.

[10] P.A.V.B. Swamy, "Linear models with random coefficients", in "Frontiers in econometrics", edited by P. Zarembka, Academic Press, New York, 1974.

[11] G.S. Maddala, "Econometrics", Mc. Graw-Hill, New York, 1977.

[12] L.R. Klein, "A textbook of econometrics", Prentice Hall, Englewood Cliffs, N.J., 1974.

[13] H. Theil, "Principles of econometrics", Wiley, New York, 1971.

[14] T.W. Anderson, "The statistical analyses of time series", Wiley, New York, 1971.

Appendix A

A vector function $f:\mathbb{R}^n \to \mathbb{R}$ is differentiable with respect to x, with gradient $\nabla_x f(x)$ (a column vector), if

$$f(x+h) = f(x) + \nabla_x f(x)'h + o(h), \quad (h \to 0). \tag{258}$$

Analogously we call a matrix function $F:M_{n,m} \to \mathbb{R}$ differentiable with respect to X with gradient $\nabla_X F(X)$ if

$$F(X+H) = F(X) + (\nabla_X F(X),H)_E + o(H), \quad (H \to 0). \tag{259}$$

Here $(\ ,\ )_E$ stands for the euclidean inner product of two matrices in $M_{n,m}$. This inner product is defined as

$$(A,B)_E := \sum_{i=1}^{n} \sum_{j=1}^{m} A_{ij} B_{ij} \qquad (A,B \in M_{n,m}). \tag{260}$$

Hence the property $\|A\|^2 = (A,A)_E$ holds.

From (259) it follows that $\nabla_X F(X)$ is a matrix in $M_{n,m}$. One can verify that the definition (259) is a simple generalization of (258).

Examples

1. $F(X) = \|X\|^2$.

$$\begin{aligned} F(X+H) &= \|X+H\|^2 = (X+H,X+H)_E = \\ &= (X,X)_E + 2(X,H)_E + (H,H)_E = \\ &= \|X\|^2 + (2X,H)_E + \|H\|^2 = \\ &= F(X) + (2X,H)_E + o(H), \qquad (H \to 0). \end{aligned}$$

Hence: $\nabla_X F(X) = 2X$.

2. $F(X) = a'Xb$, where $a \in \mathbb{R}^n$, $b \in \mathbb{R}^m$.

$$F(X+H) = a'(X+H)b = a'Xb + a'Hb = F(X) + (ab',H)_E.$$

Thus we have: $\nabla_X F(X) = ab'$.

3. $F(X) = \|Xa\|^2$, where $a \in \mathbb{R}^m$.

$$F(X+H) = \|(X+H)a\|^2 = a'(X+H)'(X+H)a =$$

$$= a'\ X'X\ a + 2a'\ X'Ha + a'H'Ha =$$

$$= \|Xa\|^2 + 2(Xa)'Ha + \|Ha\|^2 =$$

$$= F(X) + (2Xaa',H)_E + o(H), \quad (H \to 0).$$

Hence: $\nabla_X F(X) = 2Xaa'$.

Remark. In the examples 2 and 3 we used the following property:

$$x'Ay = (xy',A)_E \quad (x \in \mathbb{R}^n, y \in \mathbb{R}^m, A \in M_{n,m}).$$

Appendix B

In this section we shall prove that the following matrix is positive definite:

$$R:=\begin{pmatrix} I_n+A_0A_0'+B_0B_0' & -A_1' & & & 0 \\ -A_1 & I_n+A_1A_1'+B_1B_1' & -A_2' & & \\ & \ddots & \ddots & \ddots & \\ & & -A_{T-2} & I_n+A_{T-2}A_{T-2}'+B_{T-2}B_{T-2}' & -A_{T-1}' \\ 0 & & & -A_{T-1} & I_n+A_{T-1}A_{T-1}'+B_{T-1}B_{T-1}' \end{pmatrix}. \tag{261}$$

It's evident that $R \in M_{Tn,Tn}$. Next define for $t=0,\dots,T-1$ the matrices $R_t \in M_{(t+1)n,(t+1)n}$:

$$R_t:=\begin{pmatrix} I_n+A_0A_0'+B_0B_0' & -A_1' & & \\ -A_1 & \ddots & \ddots & \\ & \ddots & \ddots & -A_t' \\ & & -A_t & I_n+A_tA_t'+B_tB_t' \end{pmatrix}. \tag{262}$$

Then we have

$$R_{T-1} = R. \tag{263}$$

We shall use the principle of mathematical induction for proving that $R_{T-1} \in PD$ (and thus $R \in PD$). It is clear that

$$R_0 = I_n + A_0A_0' + B_0B_0' \tag{264}$$

so the property $R_0 \in PD$ holds. Now assume that for certain $t \in \{0,\dots,T-2\}$ the property $R_t \in PD$ holds. In that case we shall prove that $R_{t+1} \in PD$.

We have

$$R_{t+1} = \left[\begin{array}{cc|c} \multicolumn{2}{c|}{\multirow{2}{*}{R_t}} & 0 \\ \cline{3-3} & & -A'_{t+1} \\ \hline 0 & -A_{t+1} & I_n+A_{t+1}A'_{t+1}+B_{t+1}B'_{t+1} \end{array}\right]. \tag{265}$$

Furthermore we know that $R_{t+1} \in M_{(t+2)n,(t+2)n}$. Now let $z \in \mathbb{R}^{(t+2)n}$ and partition the vector z as follows:

$$z = \begin{bmatrix} z_1 \\ \hline z_2 \\ \hline z_3 \end{bmatrix}, \tag{266}$$

where $z_1 \in \mathbb{R}^{tn}$, $z_2, z_3 \in \mathbb{R}^n$. From (265) and (266) it follwos that

$$z'R_{t+1}z = \tilde{z}'R_t\tilde{z} - 2\, z_2'A_{t+1}'z_3 +$$

$$+ z_3'(I_n+A_{t+1}A_{t+1}'+B_{t+1}B_{t+1}')z_3. \tag{267}$$

Here we have

$$\tilde{z} := \begin{bmatrix} z_1 \\ \hline z_2 \end{bmatrix} \in \mathbb{R}^{(t+1)n}. \tag{268}$$

Hence

$$z'R_{t+1}z = z_3'(I_n+B_{t+1}B_{t+1}')z_3 + (A_{t+1}'z_3-z_2)'(A_{t+1}'z_3-z_2)$$
$$+ \tilde{z}'R_t\tilde{z} - z_2'\, z_2 \tag{269}$$

$$= z_3'(I_n+B_{t+1}B_{t+1}')z_3 + \|A_{t+1}'z_3-z_2\|^2 +$$
$$+ \tilde{z}'R_t\tilde{z}, \tag{270}$$

where the matrix $\tilde{R}_t \in M_{(t+1)n,(t+1)n}$ is defined as follows:

$$\tilde{R}_t := \left[\begin{array}{c|c|c|c}
I_n+A_0A_0'+B_0B_0' & -A_1' & & 0 \\
-A_1 & \ddots & \ddots & \\
 & \ddots & \ddots & \\
 & -A_{t-1} & I_n+A_{t-1}A_{t-1}'+B_{t-1}B_{t-1}' & -A_t' \\
0 & & -A_t & A_tA_t'+B_tB_t'
\end{array}\right]. \tag{271}$$

The property $\tilde{R}_t \in PSD$ holds because we can write:

$$\tilde{R}_t = \left[\begin{array}{c|c|c|c}
A_0 & -I_n & & 0 \\
 & A_1 & -I_n & \\
 & & \ddots & \ddots \\
 & & A_{t-1} & -I_n \\
0 & & & A_t
\end{array}\right]\left[\begin{array}{c|c|c|c}
A_0 & -I_n & & 0 \\
 & A_1 & -I_n & \\
 & & \ddots & \ddots \\
 & & A_{t-1} & -I_n \\
0 & & & A_t
\end{array}\right]' +$$

$$\begin{bmatrix} B_0 & & 0 \\ & \ddots & \\ 0 & & B_t \end{bmatrix}\begin{bmatrix} B_0 & & 0 \\ & \ddots & \\ 0 & & B_t \end{bmatrix}'. \tag{272}$$

From equation (270) it follows that

$$z'R_{t+1}z = 0 \Leftrightarrow$$

$$z_3'(I_n+B_{t+1}B'_{t+1})z_3 = 0 \wedge A'_{t+1}z_3 = z_2 \wedge \tilde{z}'\tilde{R}_t\tilde{z} = 0. \tag{273}$$

Hence $z_3 = 0$ because the matrix $I_n + B_{t+1}B'_{t+1} \in PD$. Furthermore $z_3 = 0$ implies $z_2 = 0$.
But if $z_2 = 0$ then $\tilde{z}'\tilde{R}_t\tilde{z} = \tilde{z}'R_t\tilde{z}$. So $\tilde{z} = 0$ because of the assumption $R_t \in PD$! Conclusion: $z = 0$.
We did prove now that

$$z'R_{t+1}z = 0 \Leftrightarrow z = 0 \quad . \tag{274}$$

From this equivalence if follows that $R_{t+1} \in PD$, and this completes the proof.

Appendix C

In this appendix we shall derive a sufficient condition for the positive definiteness of the matrix

$$\begin{bmatrix} R_{11} & R_{12} \\ R_{21} & R_{22} \end{bmatrix} , \tag{275}$$

which appears in the equation (146). We shall demonstrate that the matrix (275) is positive definite if one of the matrices

$$A^{(2)}A^{(2)'} , \; B^{(2)}B^{(2)'} , \; C^{(2)}C^{(2)'} , \; D^{(2)}D^{(2)'}$$

is positive definite.

We have:

$$\begin{bmatrix} \underline{r} \\ \underline{\lambda} \end{bmatrix}' \begin{bmatrix} R_{11} & R_{12} \\ R_{21} & R_{22} \end{bmatrix} \begin{bmatrix} \underline{r} \\ \underline{\lambda} \end{bmatrix} =$$

$$\begin{bmatrix} \underline{r} \\ \underline{\lambda} \end{bmatrix}' \left[\begin{array}{c|c} I_{n^{(1)}T} & 0 \\ \hline 0 & 0 \end{array} \right] \begin{bmatrix} \underline{r} \\ \underline{\lambda} \end{bmatrix} +$$

$$
+\begin{bmatrix}\underline{r}\\ \\ \underline{\lambda}\end{bmatrix}'
\left[\begin{array}{cc:cc}
B_0^{(1)} & & A_0^{(1)} & \\
 & \ddots & & \ddots \\
 & B_{T-1}^{(1)} & & A_{T-1}^{(1)} \\
\hdashline
B^{(2)} & & A^{(2)} & \\
 & \ddots & & \ddots \\
 & B^{(2)} & & A^{(2)}
\end{array}\right]
\left[\begin{array}{cc:cc}
B_0^{(1)} & & A_0^{(1)} & \\
 & \ddots & & \ddots \\
 & B_{T-1}^{(1)} & & A_{T-1}^{(1)} \\
\hdashline
B^{(2)} & & A^{(2)} & \\
 & \ddots & & \ddots \\
 & B^{(2)} & & A^{(2)}
\end{array}\right]
\begin{bmatrix}\underline{r}\\ \\ \underline{\lambda}\end{bmatrix} +
$$

$$
\begin{bmatrix}\underline{r}\\ \\ \underline{\lambda}\end{bmatrix}'
\left[\begin{array}{cc:cc}
D_0^{(1)} & & C_0^{(1)} & \\
 & \ddots & & \ddots \\
 & D^{(1)} & & C_{T-1}^{(1)} \\
\hdashline
D^{(2)} & & C^{(2)} & \\
 & \ddots & & \ddots \\
 & D^{(2)} & & C^{(2)}
\end{array}\right]
\left[\begin{array}{cc:cc}
D_0^{(1)} & & C_0^{(1)} & \\
 & \ddots & & \ddots \\
 & D_{T-1}^{(1)} & & C_{T-1}^{(1)} \\
\hdashline
D^{(2)} & & C^{(2)} & \\
 & \ddots & & \ddots \\
 & D^{(2)} & & C^{(2)}
\end{array}\right]
\begin{bmatrix}\underline{r}\\ \\ \underline{\lambda}\end{bmatrix}. \qquad (276)
$$

Hence

$$
\begin{bmatrix}\underline{r}\\ \\ \underline{\lambda}\end{bmatrix}'
\begin{bmatrix}R_{11} & R_{12}\\ \\ R_{21} & R_{22}\end{bmatrix}
\begin{bmatrix}\underline{r}\\ \\ \underline{\lambda}\end{bmatrix} = 0
$$

if and only if each of the three terms in the right-hand side of the equation (276) is equal to zero. From

$$
\begin{bmatrix}\underline{r}\\ \\ \underline{\lambda}\end{bmatrix}'
\left[\begin{array}{c:c}I_{n^{(1)}T} & 0\\ \hdashline 0 & 0\end{array}\right]
\begin{bmatrix}\underline{r}\\ \\ \underline{\lambda}\end{bmatrix} = 0
$$

it follows that $\underline{r} = \underline{0}$. If $\underline{r} = \underline{0}$ the other two conditions can be reduced to:

$$\underline{\lambda}' \begin{bmatrix} B^{(2)} \;\vdots\; A^{(2)} & & \\ & \ddots & \\ & & B^{(2)} \;\vdots\; A^{(2)} \end{bmatrix} \begin{bmatrix} B^{(2)} \;\vdots\; A^{(2)} & & \\ & \ddots & \\ & & B^{(2)} \;\vdots\; A^{(2)} \end{bmatrix}' \underline{\lambda} = 0 \quad (277)$$

and

$$\underline{\lambda}' \begin{bmatrix} D^{(2)} \;\vdots\; C^{(2)} & & \\ & \ddots & \\ & & D^{(2)} \;\vdots\; C^{(2)} \end{bmatrix} \begin{bmatrix} D^{(2)} \;\vdots\; C^{(2)} & & \\ & \ddots & \\ & & D^{(2)} \;\vdots\; C^{(2)} \end{bmatrix}' \underline{\lambda} = 0. \quad (278)$$

The equations (277) and (278) are respectively equivalent to:

$$\begin{bmatrix} \begin{array}{c} B^{(2)'} \\ \hline A^{(2)'} \end{array} & & \\ & \ddots & \\ & & \begin{array}{c} B^{(2)'} \\ \hline A^{(2)'} \end{array} \end{bmatrix} \underline{\lambda} = \underline{0} \qquad (279)$$

and

$$\begin{bmatrix} \begin{array}{c} D^{(2)'} \\ \hline C^{(2)'} \end{array} & & \\ & \ddots & \\ & & \begin{array}{c} D^{(2)'} \\ \hline C^{(2)'} \end{array} \end{bmatrix} \underline{\lambda} = \underline{0}. \qquad (280)$$

Hence

$$\begin{cases} B^{(2)'} \lambda_0 & = 0 \\ A^{(2)'} \lambda_0 + B^{(2)'} \lambda_1 & = 0 \\ \vdots & \vdots \\ A^{(2)'} \lambda_{T-2} + B^{(2)'} \lambda_{T-1} & = 0 \\ A^{(2)'} \lambda_{T-1} & = 0 \end{cases} \tag{281}$$

and

$$\begin{cases} D^{(2)'} \lambda_0 & = 0 \\ C^{(2)'} \lambda_0 + D^{(2)'} \lambda_1 & = 0 \\ \vdots & \vdots \\ C^{(2)'} \lambda_{T-2} + D^{(2)'} \lambda_{T-1} & = 0 \\ C^{(2)'} \lambda_{T-1} & = 0 \end{cases} \quad . \tag{282}$$

Now assume that e.g. $A^{(2)}A^{(2)'} \in PD$.
From (281) it follows that

$$\lambda_{T-1}' A^{(2)}A^{(2)'} \lambda_{T-1} = 0.$$

Hence

$$\lambda_{T-1} = 0.$$

Now it follows from (281) that

$$\lambda_{T-2}' A^{(2)}A^{(2)'} \lambda_{T-2} = 0,$$

hence

$$\lambda_{T-2} = 0.$$

An analogous argument results in respectively:

$$\lambda_{T-3} = 0, \ldots, \lambda_0 = 0.$$

Conclusion: $\underline{\lambda} = \underline{0}$.

The assumption that $B^{(2)}B^{(2)'}$, $C^{(2)}C^{(2)'}$ or $D^{(2)}D^{(2)'}$ is positive definite should give the same result.

We proved now that, if one of the matrices

$$A^{(2)}A^{(2)'},\ B^{(2)}B^{(2)'},\ C^{(2)}C^{(2)'},\ D^{(2)}D^{(2)'}$$

is an element of PD, the following equivalence holds:

$$\begin{bmatrix} \underline{r} \\ \underline{\lambda} \end{bmatrix}' \begin{bmatrix} R_{11} & R_{12} \\ R_{21} & R_{22} \end{bmatrix} \begin{bmatrix} \underline{r} \\ \underline{\lambda} \end{bmatrix} = 0 \Leftrightarrow \begin{bmatrix} \underline{r} \\ \underline{\lambda} \end{bmatrix} = \underline{0}.$$

This implies that under the above assumption the matrix

$$\begin{bmatrix} R_{11} & R_{12} \\ R_{21} & R_{22} \end{bmatrix}$$

is an element of PD.

Appendix D

Time series of Kleins Model I:

	C	I	W_1	Y	Π	K	G	W_2	T	tm
1920	39.8	2.7	28.8	43.7	12.7	182.8	4.6	2.2	3.4	-11
1921	41.9	-0.2	25.5	40.6	12.4	182.6	6.6	2.7	7.7	-10
1922	45.0	1.9	29.3	49.1	16.9	184.5	6.1	2.9	3.9	- 9
1923	49.2	5.2	34.1	55.4	18.4	189.7	5.7	2.9	4.7	- 8
1924	50.6	3.0	33.9	56.4	19.4	192.7	6.6	3.1	3.8	- 7
1925	52.6	5.1	35.4	58.7	20.1	197.8	6.5	3.2	5.5	- 6
1926	55.1	5.6	37.4	60.3	19.6	203.4	6.6	3.3	7.0	- 5
1927	56.2	4.2	37.9	61.3	19.8	207.6	7.6	3.6	6.7	- 4
1928	57.3	3.0	39.2	64.0	21.1	210.6	7.9	3.7	4.2	- 3
1929	57.8	5.1	41.3	67.0	21.7	215.7	8.1	4.0	4.0	- 2
1930	55.0	1.0	37.9	57.7	15.6	216.7	9.4	4.2	7.7	- 1
1931	50.9	-3.4	34.5	50.7	11.4	213.3	10.7	4.8	7.5	0
1932	45.6	-6.2	29.0	41.3	7.0	207.1	10.2	5.3	8.3	1
1933	46.5	-5.1	28.5	45.3	11.2	202.0	9.3	5.6	5.4	2
1934	48.7	-3.0	30.6	48.9	12.3	199.0	10.0	6.0	6.8	3
1935	51.3	-1.3	33.2	53.3	14.0	197.7	10.5	6.1	7.2	4
1936	57.7	2.1	36.8	61.8	17.6	199.8	10.3	7.4	8.3	5
1937	58.7	2.0	41.0	65.0	17.3	201.8	11.0	6.7	6.7	6
1938	57.5	-1.9	38.2	61.2	15.3	199.9	13.0	7.7	7.4	7
1939	61.6	1.3	41.6	68.4	19.0	201.2	14.4	7.8	8.9	8
1940	65.0	3.3	45.0	74.1	21.1	204.5	15.4	8.0	9.6	9
1941	69.7	4.9	53.3	85.3	23.5	209.4	22.3	8.5	11.6	10

All variables are measured in billions of 1934 dollars.

PREVIOUS NUMBERS

EIT 55 J. Dohmen, J. Schoeber — Approximated fixed points.

EIT 56 J.J.M. Evers — Invariant competitive equilibrium in a dynamic economy with negotiable shares.

EIT 57 W.M. van den Goorbergh — Some calculations in a three-sector model.

EIT 58 W.G. van Hulst, J.Th. van Lieshout — Investment/financial planning with endogeneous lifetimes: a heuristic approach to mixed-integer programming.

FEW 59 J.J.M. Evers, M. Schubik — A dynamic economy with shares, fiat and accounting money.

FEW 60 J.M.G. Frijns — A dynamic model of factor demand equations.

FEW 61 B.B. van der Genugten — A general approach to identification.

FEW 62 Pieter H.M. Ruys — Public goods and input-output analysis.

FEW 63 H.N. Weddepohl — Increasing returns and fixed market shares.

FEW 64 Claus Weddepohl — An equilibrium model with fixed labour time.

FEW 65 J.M.G. Frijns, A. Hempenius — Dynamic optimal factor demand under financial constraints.

FEW 66 Jack P.C. Kleijnen — Operations research and computers.

FEW 67 Jack P.C. Kleijnen, A.J. van den Burg, R.Th. van der Ham — Generalizations of simulation results: Practicality of statistical methods (part 1).

FEW 68 A.J. van den Burg, R.Th. van der Ham, J.P.C. Kleijnen — Generalizations of simulation results: Practicality of statistical methods (part 2).

FEW 69 J. Pelkmans — Customs union under increasing costs conditions

FEW 70 Anton C.M. Hopmans, Jack P.C. Kleijnen — Regression estimators in simulation.

FEW 71 J.J.M. Evers, R.T. Rockafellar, P.H.M. Ruys, H.N. Weddepohl, L.R.J. Westermann — Convex analysis and mathematical economics